

FIGURE 1

GGCATCTGCCCCGAGGAGACCACGCTCCTGGAGCTCTGCTGTCTTCTCAGGGAGACTCTGAGG
CTCTGTTGAGAATCATGCCTTTGGAGGCAGCTCATCTATTGGCAACTGCTGGCTTTGTTTTTC
CTCCCTTTTTGCCTGTGTCAAGATGAATACATGGAGTCTCCACAAACCGGAGGACTACCCCC
AGACTGCAGTAAGTGTGTCATGGAGACTACAGCTTTCGAGGCTACCAAGGCCCCCCTGGGC
CACC GGCCCTCCTGGCATTCCAGGAAACCATGGAAACAATGGCAACAATGGAGCCACTGGT
CATGAAGGAGCCAAAGGTGAGAAGGGCGACAAAGGTGACCTGGGGCCTCGAGGGGAGCGGGG
GCAGCATGGCCCCAAAGGAGAGAAGGGCTACCCGGGGATTCCACCAGAACTTCAGATTGCAT
TCATGGCTTCTCTGGCAACCCACTTCAGCAATCAGAACAGTGGGATTATCTTCAGCAGTGTT
GAGACCAACATTGGAAACTTCTTTGATGTCATGACTGGTAGATTTGGGGCCCCAGTATCAGG
TGTGTATTTCTTCACCTTCAGCATGATGAAGCATGAGGATGTTGAGGAAGTGTATGTGTACC
TTATGCACAATGGCAACACAGTCTTCAGCATGTACAGCTATGAAATGAAGGGCAAATCAGAT
ACATCCAGCAATCATGCTGTGCTGAAGCTAGCCAAAGGGGATGAGGTTTGGCTGCGAATGGG
CAATGGCGCTCTCCATGGGGACCACCAACGCTTCTCCACCTTTGCAGGATTCCTGCTCTTTG
AACTAAGTAAATATATGACTAGAATAGCTCCACTTTGGGGAAGACTTGTAGCTGAGCTGAT
TTGTTACGATCTGAGGAACATTAAAGTTGAGGGTTTTACATTGCTGTATTCAAAAAATTATT
GGTTGCAATGTTGTTACGCTACAGGTACACCAATAATGTTGGACAATTCAGGGGCTCAGAA
GAATCAACCACAAAATAGTCTTCTCAGATGACCTTGACTAATACTCAGCATCTTTATCAC
TCTTTCCTTGGCACCTAAAAGATAATTCTCCTCTGACGCAGGTTGGAAATATTTTTTTCTAT
CACAGAAGTCATTTGCAAAGAATTTTGACTACTCTGCTTTTAATTTAATACCAGTTTTTCAGG
AACCCCTGAAGTTTTTAAGTTCATTATTCTTTATAACATTTGAGAGAATCGGATGTAGTGATA
TGACAGGGCTGGGGCAAGAACAGGGGCACTAGCTGCCTTATTAGCTAATTTAGTGCCCTCCG
TGTT CAGCTTAGCCTTTGACCCTTTCCTTTTGATCCACAAAATACATTAAAACCTCTGAATTC
ACATACAATGCTATTTTAAAGTCAATAGATTTTAGCTATAAAGTGCTTGACCAGTAATGTGG
TTGTAATTTTGTGTATGTTCCCCACATCGCCCCAACTTCGGATGTGGGGTCAGGAGGTTG
AGGTTCACTATTAACAAATGTCATAAATATCTCATAGAGGTACAGTGCCAATAGATATTCAA
ATGTTGCATGTTGACCAGAGGGATTTTATATCTGAAGAACATACACTATTAATAAATACCTT
AGAGAAAGATTTTGACCTGGCTTTAGATAAAACTGTGGCAAGAAAAATGTAATGAGCAATAT
ATGGAAATAAACACACCTTTGTTAAAGATAAAAAAAA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 2

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44686

><subunit 1 of 1, 246 aa, 1 stop

><MW: 26994, pI: 6.43, NX(S/T): 0

MLWRQLIYWQLLALFFLPFCLCQDEYMESPQTGGLPPDCSKCCHGDYSFRGYQGPPGPPGPP
GIPGNHGNNGNNGATGHEGAKGEKGDLDGPRGERGQHGPKGEKGYPGIPPELQIAFMASL
ATHFSNQNSGIIIFSSVETNIGNFFDVMTGRFGAPVSGVYFFTFSMMKHEDVEEVYVYLMHNG
NTVFSMYSYEMKGKSDTSSNHAFLKLAKGDEVWLRMGNGALHGDHQRFFSTFAGFLLFETK

Important features of the protein:

Signal peptide:

amino acids 1-22

Motif name: Clq domain signature.

amino acids 137-167

Clq domain proteins.

amino acids 135-169, 202-221, 235-244, 57-91, 60-94, 54-88, 81-114, 78-111, 63-96, 51-84, 45-78, 48-81, 33-66, 66-99 and 42-75

FIGURE 3

GAGAGAATAGCTACAGATTCTCCATCCTCAGTCTTTGCAAGGCGACAGCTGTGCCAGCCGGG
CTCTGGCAGGCTCCTGGCAGCATGGCAGTGAAGCTTGGGACCCTCCTGCTGGCCCTTGCCCT
GGGCCTGGCCCAGCCAGCCTCTGCCCCGCCGGAAGCTGCTGGTGTTTCTGCTGGATGGTTTTT
GCTCAGACTACATCAGTGATGAGGCGCTGGAGTCATTGCCTGGTTTTCAAAGAGATTGTGAGC
AGGGGAGTAAAAGTGGATTACTTGACTCCAGACTTCCCTAGTCTCTCGTATCCCAATTATTA
TACCCTAATGACTGGCCGCCATTGTGAAGTCCATCAGATGATCGGGAACTACATGTGGGACC
CCACCACCAACAAGTCCTTTGACATTGGCGTCAACAAAGACAGCCTAATGCCTCTCTGGTGG
AATGGATCAGAACCTCTGTGGGTCACTCTGACCAAGGCCAAAAGGAAGGTCTACATGTACTA
CTGGCCAGGCTGTGAGGTTGAGATTCTGGGTGTCAGACCCACCTACTGCCTAGAATATAAAA
ATGTCCCAACGGATATCAATTTTGCCAATGCAGTCAGCGATGCTCTTGACTCCTTCAAGAGT
GGCCGGGCCGACCTGGCAGCCATATACCATGAGCGCATTGACGTGGAAGGCCACCACTACGG
GCCTGCATCTCCGCAGAGGAAAGATGCCCTCAAGGCTGTAGACACTGTCCTGAAGTACATGA
CCAAGTGGATCCAGGAGCGGGGCCTGCAGGACCGCCTGAACGTCATTATTTTCTCGGATCAC
GGAATGACCGACATTTTCTGGATGGACAAAGTGATTGAGCTGAATAAGTACATCAGCCTGAA
TGACCTGCAGCAAGTGAAGGACCGCGGGCCTGTTGTGAGCCTTTGGCCGGCCCCCTGGGAAAC
ACTCTGAGATATATAACAACTGAGCACAGTGGAACACATGACTGTCTACGAGAAAGAAGCC
ATCCCAAGCAGGTTCTATTACAAGAAAGGAAAGTTTGTCTCTCCTTTGACTTTAGTGGCTGA
TGAAGGCTGGTTCATAACTGAGAATCGAGAGATGCTTCCGTTTTGGATGAACAGCACCGGCA
GGCGGGAAGGTTGGCAGCGTGGATGGCACGGCTACGACAACGAGCTCATGGACATGCGGGGC
ATCTTCCTGGCCTTCGGACCTGATTTCAAATCCAACCTTCAGAGCTGCTCCTATCAGGTCGGT
GGACGTCTACAATGTCATGTGCAATGTGGTGGGCATCACCCCGCTGCCCAACAACGGATCCT
GGTCCAGGGTGATGTGCATGCTGAAGGGCCGCGCCGGCACTGCCCCGCCTGTCTGGCCCAGC
CACTGTGCCCTGGCACTGATTCTTCTCTTCCTGCTTGCATAACTGATCATATTGCTTGTCTC
AGAAAAAACACCATCAGCAAAGTGGGCCTCCAAGCCAGATGATTTTCATTTTATGTGTGA
ATAATAGCTTCATTAACACAATCAAGACCATGCACATTGTAAATACATTATTCTTGGATAAT
TCTATACATAAAAGTTCCTACTTGTTAAA

FIGURE 4

MAVKLGTLALLALGLAQPASARRKLLVFLLDGFRSDYISDEALES LPGFKEIVSRGVKVDY
LTPDFPSLSYPNYYTLMTGRHCEVHQMIGNYMWDPTTNKSFDIGVNKDSL MPLWWNGSEPLW
VTLTAKARKVYMYYPGCEVEILGVRPTYCLEYKNVPTDINFANAVSDALDSFKSGRADLAA
IYHERIDVEGHHYGPASPQRKDALKAVDTVLKYMTKWIQERGLQDRNLNVIIFSDHGMTDIFW
MDKVIELNKYISLNDLQQVKDRGPVVSLWPAPGKHSEIYNKLSTVEHMTVYEKEAIPSRFYY
KKGK FVSPLTLVADEGWFITENREMLPFWMNSTGRREGWQRGWHGYDNELMDMRGIFLAFGP
DFKSNFRAAPIRSVDVYNVMCNVVGITPLPNNGSWSRVMCMLKGRAGTAPPVWPSHCALALI
LLFLLA

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation sites.

amino acids 100-104, 118-122, 341-345, 404-408

N-myristoylation sites.

amino acids 148-154, 365-371

Amidation site.

amino acids 343-347

FIGURE 5

GCCAGGTGTGCAGGCCGCTCCAAGCCCAGCCTGCCCCGCTGCCGCCACC**ATG**ACGCTCCTCC
CCGGCCTCCTGTTTCTGACCTGGCTGCACACATGCCTGGCCCACCATGACCCCTCCCTCAGG
GGGCACCCCCACAGTCACGGTACCCACACTGCTACTCGGCTGAGGAACTGCCCCCTCGGCCA
GGCCCCCCCCACACCTGCTGGCTCGAGGTGCCAAGTGGGGGCAGGCTTTGCCTGTAGCCCTGG
TGTCCAGCCTGGAGGCAGCAAGCCACAGGGGGAGGCACGAGAGGCCCTCAGCTACGACCCAG
TGCCCGGTGCTGCGGCCGGAGGAGGTGTTGGAGGCAGACACCCACCAGCGCTCCATCTCACC
CTGGAGATACCGTGTGGACACGGATGAGGACCGCTATCCACAGAAGCTGGCCTTCGCCGAGT
GCCTGTGCAGAGGCTGTATCGATGCACGGACGGGCCGCGAGACAGCTGCGCTCAACTCCGTG
CGGCTGCTCCAGAGCCTGCTGGTGCTGCGCCGCCGGCCCTGCTCCCGCGACGGCTCGGGGCT
CCCCACACCTGGGGCCTTTGCCTTCCACACCGAGTTCATCCACGTCCCCGTCGGCTGCACCT
GCGTGCTGCCCCGTTCAGTG**TGA**ACCGCCGAGGCCGTGGGGCCCCTAGACTGGACACGTGTGC
TCCCCAGAGGGCACCCCCTATTTATGTGTATTTATTGTTATTTATATGCCTCCCCCAACACT
ACCCTTGGGGTCTGGGCATTCCCCGTGTCTGGAGGACAGCCCCCACTGTTCTCCTCATCTC
CAGCCTCAGTAGTTGGGGGTAGAAGGAGCTCAGCACCTCTTCCAGCCCTTAAAGCTGCAGAA
AAGGTGTCACACGGCTGCCTGTACCTTGGCTCCCTGTCCTGCTCCCGGCTTCCCTTACCCTA
TCACTGGCCTCAGGCCCCGCAGGCTGCCTCTTCCCAACCTCCTTGGAAGTACCCCTGTTTCT
TAAACAATTATTTAAGTGTACGTGTATTATTAACTGATGAACACATCCCCAAAA

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 6

MTLLPGLLFLTWLHTCLAHHDPSLRGHPHSHGTPHCYSAEELPLGQAPPHLLARGAKWGQAL
PVALVSSLEAASHRGRHERPSATTQCPVLRPEEVLEADTHQRSISPWRYRVDTDED RYPQKL
AFAECLCRGCIDARTGRETAALNSVRLQLSLLVLRRRPCSRDGSGLPTPGAFAFHTEFIHVP
VGCTCVLPRSV

Important features:

Signal peptide:

amino acids 1-18

Tyrosine kinase phosphorylation site.

amino acids 112-121

N-myristoylation sites.

amino acids 32-38, 55-61, 133-139

Leucine zipper pattern.

amino acids 3-25

Homologous region to IL-17.

amino acids 99-195

FIGURE 7

CGGCCAGGGCGCCGACAGCCCGACCTCACCAGGAGAAC**ATG**CAGCTCGGCACTGGGCTCCTG
CTGGCCGCCGTCTGTAGCCTGCAGCTGGCTGCAGCCGAAGCCATATGGTGTCAACAGTGCAC
GGGCTTCGGAGGGTGCTCCCATGGATCCAGATGCCTGAGGGACTCCACCCACTGTGTACCA
CTGCCACCCGGGTCTCAGCAACACCGAGGATTTGCCTCTGGTCACCAAGATGTGCCACATA
GGCTGCCCCGATATCCCCAGCCTGGGCCTGGGCCCCCTACGTATCCATCGCTTGCTGCCAGAC
CAGCCTCTGCAACCATGAC**TGA**CGGCTGCCCTCCTCCAGGCCCCCGGACGCTCAGCCCCAC
AGCCCCACAGCCTGGCGCCAGGGCTCACGGCCGCCCTCCCTCGAGACTGGCCAGCCCACC
TCTCCCGGCCTCTGCAGCCACCGTCCAGCACCGCTTGTCTAGGGAAGTCCTGCGTGGAGTC
TTGCCTCAATCTGCTGCCGTCCAAGCCTGGGGCCCATCGTGCCTGCCGCCCCCTTCAGGTCCC
GACCTCCCCACAATAAAATGTGATTGGATCGTGTGGTACAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 8

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77623

><subunit 1 of 1, 97 aa, 1 stop

><MW: 10160, pI: 6.56, NX(S/T): 0

MLGTTGLLLAAVLSLQLAAAEAIWCHQCTGFGGCSHGSRCLRDSTHCVTTATRVLSNTEDLP
LVTKMCHIGCPDIPSLGLGPYVSIACCQTS�CNHD

Important features of the protein:

Signal peptide:

amino acids 1-20

N-myristoylation sites.

amino acids 6-11 and 33-38

Prokaryotic membrane lipoprotein lipid attachment sites.

amino acids 24-34 and 78-88

FIGURE 9

CCAGGACCAGGGCGCACCGGCTCAGCCTCTCACTTGTCAGAGGCCGGGGAAGAGAAGCAAAG
CGCAACGGTGTGGTCCAAGCCGGGGCTTCTGCTTCGCCTCTAGGACATACACGGGACCCCT
AACTTCAGTCCCCCAAACGCGCACCCCTCGAAGTCTTGAAGTCCAGCCCCGCACATCCACGCG
CGGCACAGGCGCGGCAGGCGGCAGGTCCCGGCCGAAGGCGATGCGCGCAGGGGGTTCGGGCAG
CTGGGCTCGGGCGGGCGGGAGTAGGGCCCCGGCAGGGAGGCAGGGAGGCTGCATATTCAGAGTC
GCGGGCTGCGCCCTGGGCAGAGGCCGCCCTCGCTCCACGCAACACCTGCTGCTGCCACCGCG
CCGCGATGAGCCGCGTGGTCTCGCTGCTGCTGGGCGCCGCGCTGCTCTGCGGCCACGGAGCC
TTCTGCGCGCCGCGTGGTCAGCGGCCAAAAGGTGTGTTTTGCTGACTTCAAGCATCCCTGCTA
CAAAATGGCCTACTTCCATGAACTGTCCAGCCGAGTGAGCTTTCAGGAGGCACGCCTGGCTT
GTGAGAGTGAGGGAGGAGTCCTCCTCAGCCTTGAGAATGAAGCAGAACAGAAGTTAATAGAG
AGCATGTTGCAAAACCTGACAAAACCCGGGACAGGGATTTCTGATGGTGATTTCTGGATAGG
GCTTTGGAGGAATGGAGATGGGCAAACATCTGGTGCCTGCCAGATCTCTACCAGTGGTCTG
ATGGAAGCAATTCAGTACCGAACTGGTACACAGATGAACCTTCCTGCGGAAGTGAAAAG
TGTGTTGTGATGTATCACCAACCAACTGCCAATCCTGGCCTTGGGGGTCCCTACCTTTACCA
GTGGAATGATGACAGGTGTAACATGAAGCACAATTATATTTGCAAGTATGAACCAGAGATTA
ATCCAACAGCCCCTGTAGAAAAGCCTTATCTTACAAATCAACCAGGAGACACCCATCAGAAT
GTGGTTGTTACTGAAGCAGGTATAATTCCCAATCTAATTTATGTTGTTATACCAACAATACC
CCTGCTCTTACTGATACTGGTTGCTTTTGGAACTGTTGTTTCCAGATGCTGCATAAAAGTA
AAGGAAGAACAAAACTAGTCCAAACCAGTCTACACTGTGGATTTCAAAGAGTACCAGAAAA
GAAAGTGGCATGGAAGTATAATAACTCATTGACTTGGTTCCAGAATTTTGTAAATCTGGATC
TGATAAGGAATGGCATCAGAACAAATAGCTTGGAAATGGCTTGAATCACAAAGGATCTGCAA
GATGAACTGTAAGCTCCCCCTTGAGGCAAATATTAAGTAATTTTTATATGTCTATTATTTTC
ATTTAAAGAATATGCTGTGCTAATAATGGAGTGAGACATGCTTATTTTGCTAAAGGATGCAC
CCAACTTCAAACCTCAAGCAAATGAAATGGACAATGCAGATAAAGTTGTTATCAACACGTC
GGGAGTATGTGTGTAGAAAGCAATTCCTTTTATTTCTTTACCTTTTATAAGTTGTTATCTA
GTCAATGTAATGTATATTGTATTGAAATTTACAGTGTGCAAAAGTATTTTACCTTTGCATAA
GTGTTTGATAAAAAATGAACTGTTCTAATATTTATTTTATGGCATCTCATTTTTCAATACAT
GCTCTTTTGATTAAAGAACTTATTACTGTTGTCAACTGAATTCACACACACACAAATATAG
TACCATAGAAAAAGTTTGTTTTCTCGAAATAATTCATCTTTTACGCTTCTCTGCTTTTGGTCA
ATGTCTAGGAAATCTCTTCAGAAATAAGAAGCTATTTTCAATTAAGTGATATAAACCTCCTC
AAACATTTTACTTAGAGGCAAGGATTGTCTAATTTCAATTTGTGCAAGACATGTGCCTTATAA
TTATTTTATAGCTTAAAATTAAACAGATTTTGTAAATAATGTAACCTTGTTAATAGGTGCATAA
ACACTAATGCAGTCAATTTGAACAAAAGAAGTGACATACACAATATAAATCATATGTCTTCA
CACGTTGCCATATAATGAGAAGCAGCTCTCTGAGGGTTCTGAAATCAATGTGGTCCCTCTC
TTGCCCACTAAACAAAGATGGTTGTTTCGGGGTTTGGGATTGACACTGGAGGCAGATAGTTGC
AAAGTTAGTCTAAGGTTTCCCTAGCTGTATTTAGCCTCTGACTATATTAGTATACAAAGAGG
TCATGTGGTTGAGACCAGGTGAATAGTCACTATCAGTGTGGAGACAAGCACAGCACACAGAC
ATTTTAGGAAGGAAAGGAACTACGAAATCGTGTGAAATGGGTTGGAACCCATCAGTGATCG
CATATTCATTGATGAGGGTTTGTCTTGAGATAGAAAATGGTGGCTCCTTTCTGTCTTATCTCC
TAGTTTCTTCAATGCTTACGCCTTGTCTTCTCAAGAGAAAGTTGTAACCTCTCTGGTCTTCA
TATGTCCCTGTGCTCCTTTTAACCAAATAAAGAGTTCTTGTCTTCTGGGGGAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 10

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA79230

><subunit 1 of 1, 273 aa, 1 stop

><MW: 30431, pI: 6.79, NX(S/T): 3

MSRVVSLLLGAALLCGHGAFRRVVSQGKVCFADEFKHPCKMAYFHELSSRVSFQEARLACE
SEGGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDFWIGLWRNGDGQTSGACPDLYQWSDG
SNSQYRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNICKYEPEINP
TAPVEKPYLTNQPGDTHQNVVVTEAGIIPNLIYVVIPTIPLLLLILVAFGTCCFQMLHKSKG
RTKTSPNQSTLWISKSTRKESGMEV

Important features of the protein:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 214-235

N-glycosylation sites.

amino acids 86-89 and 255-258

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 266-269

N-myristoylation sites.

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-145
and 212-217

FIGURE 11

GGAGAAATGGAGAGAGCAGTGAGAGTGGAGTCCGGGGTCCCTGGTCCGGGTGGTCTGTCTGCTCCTGGCATGCCCTG
CCACAGCCACTGGGCCCCGAAGTTGCTCAGCCTGAAGTAGACACCACCCTGGGTTCGTGTGCGAGGCCGGCAGGTGG
GCGTGAAGGGCACAGACCGCTTGTGAATGTCTTTCTGGGCATTCCATTTGCCAGCCGCCACTGGGCCCTGACC
GGTCTCAGCCCCACACCCAGCACAGCCCTGGGAGGGTGTGCGGGATGCCAGCACTGCGCCCCCAATGTGCCTAC
AAGACGTGGAGAGCATGAACAGCAGCAGATTTGTCTCAACGGAAAACAGCAGATCTTCTCCGTTTCAGAGGACT
GCCTGGTCCCTCAACGTCTATAGCCCAGCTGAGGTCCCCGAGGGTCCGGTAGGCCGGTTCATGGTATGGGTCCATG
GAGGCGCTCTGATAACTGGCGCTGCCACCTCCTACGATGGATCAGCTCTGGCTGCCTATGGGGATGTGGTCTGTGG
TTACAGTCCAGTACCGCCTTGGGGTCTTGGCTTCTTACGACTGGAGATGAGCATGCACCTGGCAACCAGGGCT
TCCTAGATGTGGTAGCTGCTTTGCGCTGGGTGCAAGAAAACATCGCCCCCTTCGGGGGTGACCTCAACTGTGTCA
CTGTCTTTGGTGGATCTGCCGGTGGGAGCATCATCTCTGGCCTGGTCTGTCCCCAGTGGCTGCAGGGCTGTTCC
ACAGAGCCATCACACAGAGTGGGGTTCATCACCACCCAGGGATCATCGACTCTCACCTTGGCCCCCTAGCTCAGA
AAATCGCAAAACACCTTGGCCTGCAGCTCCAGCTCCCCGGCTGAGATGGTGCAGTGCCTTCAGCAGAAAGAAGGAG
AAGAGCTGGTCTTGAAGAAGCTGAAAAATACTATCTATCCTCTCACCGTTGATGGCACTGTCTTCCCCAAAA
GCCCCAAGGAACCTCTGAAGGAGAAGCCCTTCCACTCTGTGCCCTTCCCTCATGGGTGTCAACAACCATGAGTTCA
GCTGGCTCATCCCCAGGGGCTGGGGTCTCCTGGATACAATGGAGCAGATGAGCCGGGAGGACATGCTGGCCATCT
CAACACCCGTCTTGACCACTGTGGATGTGCCCCCTGAGATGATGCCACCCGTATAGATGAATACCTAGGAAGCA
ACTCGGACGCACAAGCCAAATGCCAGGCGTTCAGGAATTCATGGGTGACGTATTTCATCAATGTTCCACCGTCA
GTTTTTCAAGATACCTTCGAGATTCTGGAAGCCCTGTCTTTTTCTATGAGTTCAGCATCGACCCAGTTCTTTTG
CGAAGATCAAACCTGCCTGGGTGAAGGCTGATCATGGGGCCGAGGGTCTTTTGTGTTCCGAGGTCCCTTCTCA
TGGACGAGAGCTCCCGCCTGGCCTTTCAGAGGCCACAGAGGAGGAGAAGCAGCTAAGCCTCACCATGATGGCCC
AGTGGACCCACTTTGCCCGGACAGGGGACCCCAATAGCAAGGCTCTGCCTCCTTGGCCCCAATTCAACAGGGCGG
AACAATATCTGGAGATCAACCCAGTGCCACGGGCGGACAGAAGTTCAGGGAGGCCCTGGATGCAGTTCTGGTCAG
AGACGCTCCCCAGCAAGATACAACAGTGGCACCAGAAGCAGAAGAACAGGAAGGCCACAGGAGACCTCTGAGGCC
AGGCGTGAACCTTCTTGGCTGGGGCAAACCACTCTTCAAGTGGTGGCAGAGTCCAGCACGGCAGCCCCCTCTC
CCCCTGCTGAGACTTTAATCTCCACAGCCCTTAAAGTGTGCGCCGCTCTGTGACTGGAGTTATGCTCTTTTGAA
ATGTCACAAGGCCGCTCCACCTCTGGGGCATTGTACAAGTTCTTCCCTCTCCCTGAAGTGCCTTTCTGCTTT
CTTCGTGGTAGGTTCTAGCACATTCCTCTAGCTTCTGGAGGACTCACTCCCCAGGAAGCCTTCCCTGCCTTCTC
TGGGCTGTGCGGCCCCGAGTCTGCGTCCATTAGAGCACAGTCCACCCGAGGCTAGCACCGTGTCTGTGTCTGTCT
CCCCCTCAGAGGAGCTCTCTCAAAATGGGGATTAGCCTAACCCCACTCTGTACCCACACAGGATCGGGTGGGA
GCTGGAGCTAGGGGTGTTTGTGAGTGAGTGAGTGAACACAGAATATGGGAATGGCAGCTGCTGAACTTGAAC
CCAGAGCCTTCAGGTGCCAAAGCCATACTCAGGCCCCACCCGACATTTGCCACCCTGGCCAGAAGGGTGCATGCC
AATGGCAGAGACCTGGGATGGGAGAAGTCTGGGGCGCCAGGGGATCCAGCCTAGAGCAGACCTTAGCCCCCTGAC
TAAGGCCTCAGACTAGGGCGGGAGGGGTCTCCTCCTCTCTGCTGCCAGTCTGGCCCCCTGCACAAGACAACAGA
ATCCATCAGGGCCATGAGTGTACCCAGACCTGACCCTCACCAATTCAGCCCCCTGACCCTCAGGACGCTGGATG
CCAGCTCCCAGCCCCAGTGCCGGGTCTCCTCCTCCTTCTGGCTTGGGGAGACCAGTTTCTGGGGAGCTTCCAAG
AGCACCCACCAAGACACAGCAGGACAGGCCAGGGGAGGGCATCTGGACCAGGGCATCCGTGCGGCTATTGTACA
GAGAAAAGAAGAGACCCACCCACTCGGGCTGCAAAAGGTGAAAAGCACCAAGAGGTTTTTCAGATGGAAGTGAGAG
GTGACAGTGTGCTGGCAGCCCTCACAGCCCTCGCTTGTCTCCTGCCGCTCTGCCTGGGCTCCCACTTTGGCA
GCACTTGAGGAGCCCTTCAACCCGCCGCTGCACTGTAGGAGCCCTTTCTGGGCTGGCCAAGGCCGGAGCCAGCT
CCCTCAGCTTGGGGGAGGTGCGGAGGGAGAGGGGCGGGCAGGAACCGGGGCTGCGCGCAGCGCTTGGGGCCAG
AGTGAGTTCCGGGTGGGCGTGGGCTCGGCGGGGCCCCACTCAGAGCAGCTGGCCGGCCCCAGGAGTGAGGGCCT
TAGCACCTGGGCCAGCAGCTGCTGTGCTCGATTTCTCGCTGGGCCCTAGCTGCCTCCCCGCGGGGAGGGCTCGG
GACCTGCAGCCCTCCATGCCTGACCCTCCCCCACCCCCCGTGGGCTCCTGTGCGGCCGAGCCTCCCCAAGGAG
CGCCGCCCCCTGCTCCACAGCGCCAGTCCCATCGACCACCAAGGGCTGAGGAGTGCGGGTGCACAGCGCGGGA
CTGGCAGGCAGCTCCACCTGCTGCCCCAGTGTGGATCCACTGGGTGAAGCCAGCTGGGCTCCTGAGTCTGGTGG
GGACTTGGAGAACCTTTATGTCTAGCTAAGGGATTGTAAATACACCGATGGGCACTCTGTATCTAGCTCAAGGTT
TGTAACACACCAATCAGCACCCCTGTGTCTAGCTCAGTGTGTGTGAATGCACCAATCCACACTCTGTATCTGGCT
ACTCTGGTGGGACTTGGAGAACCCTTTGTGTCCACACTCTGTATCTAGCTAATCTAGTGGGGATGTGGAGAACCCT
TTGTGTCTAGCTCAGGGATCGTAAACGCACCAATCAGCACCCCTGTCAAAACAGACCCTTGACTCTCTGTAAAT
GGACCAATCAGCAGGATGTGGGTGGGGCGAGACAAGAGAATAAAAGCAGGCTGCCTGAGCCAGCAGTGACAACCC
CCCTCGGGTCCCCCTCCACGCCGTGGAAGCTTTGTCTTTGCTCTTTGCAATAAATCTTGCTACTGCCAAAA

FIGURE 12

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA79862

><subunit 1 of 1, 571 aa, 1 stop

><MW: 62282, pI: 5.56, NX(S/T): 1

MERAVRVESGVLVGVVCLLLACPATATGPEVAQPEVDTTLGRVRGRQVGVKGTDRLVNVFLG
IPFAQPPLGPDRFSAPHPAQWEGVRDASTAPPMCLQDVESMNSSRFVLNGKQQIFSVSEDC
LVLNVYSPA EVPAGSGRPVMVWVHGGALITGAATSYDGSALAA YGDVVVVTVQYRLGVLGFF
STGDEHAPGNQGFLDVVAALRWVQENIAPFGGDLNCVTVFEGGSAGGSIIISGLVLS PVAAGLF
HRAITQSGVITTPGIIDSHPWPLAQKIAN TLACSSSSPAEMVQCLQQKEGEELVLSKKLKNT
IYPLTVDGTVPFKSPKELLKEKPFH SVPFMLGMVNNHEFSWLI PRGWGLLDTMEQMSREDMLA
ISTPVLTSLDVPPPEMMPTVIDEYLGSNSDAQAKCQAFQEFMGDVFINVPTVSFSRYLRDSGS
PVFFYEFQHRPSSFAKIKPAWVKADHGAEGAFVFGGPFLMDESSRLAFPEATEEEKQLSLTM
MAQWTHFARTGDPNSKALPPWPQFNQAEQYLEINPVPRAGQKFREAWMQFWSETLPSKIQQW
HQKQKNRKAQEDL

Important features of the protein:

Signal peptide:

amino acids 1-27

Transmembrane domain:

amino acids 226-245

N-glycosylation site.

amino acids 105-109

N-myristoylation sites.

amino acids 10-16, 49-55, 62-68, 86-92, 150-156, 155-161,
162-168, 217-223, 227-233, 228-234, 232-238, 262-268, 357-363,
461-467

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 12-23

Carboxylesterases type-B serine active site.

amino acids 216-232

FIGURE 13

CATGGAGCCTCTTGCAGCTTACCCGCTAAAATGTTCCGGGCCCAGAGCAAAGGTATTTGCAG
TTTTGCTGTCTATAGTTCTATGCACAGTAACGCTATTTCTTCTACAACATAAAATTCCTCAAA
CCTAAAATCAACAGCTTTTATGCCTTTGAAGTGAAGGATGCAAAAGGAAGAACTGTTTCTCT
GGAAAAGTATAAAGGCCAAAGTTTCACTAGTTGTAAACGTGGCCAGTGACTGCCAACTCACAG
ACAGAAATTACTTAGGGCTGAAGGAAGTGCACAAAGAGTTTGGACCATCCCACCTCAGCGTG
TTGGCTTTTCCCTGCAATCAGTTTGGAGAATCGGAGCCCCGCCCAAGCAAGGAAGTAGAATC
TTTTGCAAGAAAAAACTACGGAGTAACTTTCCCCATCTTCCACAAGATTAAGATTCTAGGAT
CTGAAGGAGAACCTGCATTTAGATTTCTTGTGATTCTTCAAAGAAGGAACCAAGGTGGAAT
TTTTGGAAGTATCTTGTCAACCCTGAGGGTCAAGTTGTGAAGTTCTGGAGGCCAGAGGAGCC
CATTGAAGTCATCAGGCCTGACATAGCAGCTCTGGTTAGACAAGTGATCATAAAAAAGAAAG
AGGATCTA**TGA**GAAATGCCATTGCGTTTCTAATAGAACAGAGAAATGTCTCCATGAGGGTTTG
GTCTCATTTTTAAACATTTTTTTTTTTGGAGACAGTGTCTCACTCTGTCACCCAGGCTGGAGTG
CAGTAGTGC GTTCTCAGCTCATTGCAACCTCTGCCTTTTTTAAACATGCTATTAAATGTGGCA
ATGAAGGATTTTTTTTTTAATGTTATCTTGCTATTAAGTGGAATGAATGTTCCCAGGATGAG
GATGTTACCCAAAGCAAAAATCAAGAGTAGCCAAAGAATCAACATGAAATATATTAAGTACT
TCCTCTGACCATACTAAAGAATTCAGAATACACAGTGACCAATGTGCCTCAATATCTTATTG
TTCAACTTGACATTTTCTAGGACTGTACTTGATGAAAATGCCAACACACTAGACCACTCTTT
GGATTCAAGAGCACTGTGTATGACTGAAATTTCTGGAATAACTGTAAATGGTTATGTTAATG
GAATAAAACACAAATGTTGAAAAATGTAAAATATATATACATAGATTCAAATCCTTATATAT
GTATGCTTGTTTTGTGTACAGGATTTTGTTTTTTCTTTTTTAAGTACAGGTTCCTAGTGTTTT
ACTATAACTGTCACTATGTATGTAAGTACATATATAAATAGTCATTTATAAATGACCGTAT
TATAACATTTGAAAAAGTCTTCATCAAAAAAAAAAAAAA

FIGURE 14

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA80136

><subunit 1 of 1, 209 aa, 1 stop

><MW: 23909, pI: 9.68, NX(S/T): 0

MEPLAAYPLKCSGPRAKVFAVLLSIVLCTVTLELLQLKFLKPKINSFYAFEVKDAKGRTVSL
EKYKGVSLVVNVASDCQLTDRNYLGLKELHKEFGPSHFVLAFFPCNQFGESEPRPSKEVES
FARKNYGVTFPIFHKIKILGSEGEPAFRFLVDSSKKEPRWNFWKYLVNPEGQVVKFWRPEEP
IEVIRPDIAALVRQVVIKKKEDL

Important features of the protein:

Signal peptide:

amino acids 1-31

Glutathione peroxidases signature 2.

amino acids 104-112

Glutathione peroxidases.

amino acids 57-82

FIGURE 15

TGTCGCCTGGCCCTCGCCATGCAGACCCCGCGAGCGTCCCCTCCCCGCCCCGGCCCTCCTGCTTCTGCTGCTGCTA
CTGGGGGGGCGCCACGGCCCTCTTCTGAGGAGCCGCCGCGCTTAGCGTGGCCCCCAGGGACTACCTGAACCAC
TATCCCGTGTGTTGTGGGCGAGCGGGCCCCGACGCGCTGACCCCGCAGAAGGTGCTGACGACCTCAACATCCAGCGA
GTCCTGCGGGTCAACAGGACGCTGTTTATGTTGGGACAGGGACAACCTCTACCGCGTAGAGCTGGAGCCCCCAGC
TCCACGGAGCTGCGGTACAGAGGAAGCTGACCTGGAGATCTAACCCAGCGACATAAACGTGTGTCGGATGAAG
GGCAAACAGGAGGGCGAGTGTGAAACTTCGTAAAGGTGCTGCTCCTTCGGGACGAGTCCACGCTCTTTGTGTGC
GGTTCCAACGCCTTCAACCCGGTGTGCGCAACTACAGCATAGACACCCTGCAGCCCGTCGGAGACAACATCAGC
GGTATGGCCCGCTGCCCCGTACGACCCCAAGCAGCÇAATGTTGCCCTCTTCTCTGACGGGATGCTCTTCACAGCT
ACTGTTACCGACTTCTAGCCATTGATGCTGTCATCTACCGCAGCCTCGGGGACAGGCCACCCTGCGCACCGTG
AAACATGACTCCAAGTGGTTCAAAGAGCCTTACTTTGTCCATGCGGTGGAGTGGGGCAGCCATGTCTACTTCTTC
TTCCGGGAGATTGCGATTGAGTTTAACTACCTGGAGAAGGTGGTGGTGTCCCGCGTGGCCCCGAGTGTGCAAGAAC
GACGTGGGAGGCTCCCCCGCGTGTGGAGAAAGCAGTGGACGTCCTTCTGAAGGCGCGGCTCAACTGCTCTGTA
CCCCGAGACTCCCATTTCTACTTCAACGTGCTGCAGGCTGTACGGGCGTGGTCAGCCTCGGGGGCGGGCCCGTG
GTCCTGGCCGTTTTTTTCCACGCCCAGCAACAGCATCCCTGGCTCGGCTGTCTGCGCCTTTGACCTGACACAGGTG
GCAGCTGTGTTTGAAGGCCGCTTCCGAGAGCAGAAGTCCCCCGAGTCCATCTGGACGCCGGTGCCGGAGGATCAG
GTGCCTCGACCCCGGCCCCGGGTGCTGCGCAGCCCCCGGGATGCAGTACAATGCCCTCCAGCGCCTTGCCGGATGAC
ATCCTCAACTTTGTCAAGACCCACCCTCTGATGGACGAGGCGGTGCCCTCGCTGGGCCATGCGCCCTGGATCCTG
CGGACCCTGATGAGGCACCAGCTGACTCGAGTGGCTGTGGACGTGGGAGCCGGCCCCCTGGGGCAACCAGACCGTT
GTCTTCTGGGTTCTGAGGCGGGGACGGTCTCAAGTTCTCGTCCGGCCCAATGCCAGCACCTCAGGGACGCTCT
GGGCTCAGTGTCTTCTGGAGGAGTTTGAGACCTACCGGCCGACAGGTGTGGACGGCCCGCGGTGGCGAGACA
GGGCGAGCGGCTGCTGAGCTTGGAGCTGGACGCAGCTTCCGGGGGCGCTGCTGGCTGCCTTCCCCCGCTGCGTGGTC
CGAGTGCCTGTGGCTCGCTGCCAGCAGTACTCGGGGTGTATGAAGAACTGTATCGGCAGTCAGGACCCCTACTGC
GGGTGGGCCCCGACGGCTCTGTCATCTTCTCAGCCCGGGCACCAGAGCCGCTTTGAGCAGGACGTGTCCGGG
GCCAGCACCTCAGGCTTAGGGGACTGCACAGGACTCCTGCGGGCCAGCCTCTCCGAGGACCGCGCGGGGCTGGTG
TCGGTGAACCTGCTGGTAACGTGCTCGGTGGCGGCCCTTCGTGGTGGGAGCCGTGGTGTCCGGCTTCAGCGTGGGC
TGGTTCTGTTGGCCCTCCGTGAGCGGCGGGAGCTGGCCCGGCGCAAGGACAAGGAGGCCATCCTGGCGCACGGGGCG
GGCGAGGCGGTGCTGAGCGTCAGCCGCTGGGCGAGCGCAGGGCGCAGGGTCCCGGGGGCGGGGGCGGAGGCGGT
GGCGGTGGCGCCGGGGTTCCCCCGAGGCCCTGCTGGCGCCCTGATGCAGAACGGCTGGGCCAAGGCCACGCTG
CTGCAGGGCGGGCCCCACGACCTGGACTCGGGGTGCTGCCCCAGCCCGAGCAGACGCCGCTGCCGAGAACGCG
CTGCCCCACTCCGACCCCGACCCCGACGCCCTGGGCCCGCGCCTGGGACCACGGCCACCCCTGCTCCCGGCC
TCCGCTTCATCCTCCCTCCTGCTGCTGGCGCCCGCCGGGCCCCCGAGCAGCCCCCGCGCCTGGGGAGCCGACC
CCCGACGGCCGCTCTATGCTGCCCGGCCCGCGCGCCTCCACGGCGACTTCCCGCTCACCCCCACGCCAGC
CCGGACCGCCGGCGGGTGGTGTCCGCGCCACGGGCCCTTGGACCCAGCCTCAGCCGCCGATGGCCTCCCGCGG
CCCTGGAGCCCCGCCCCGACGGGCAGCCTGAGGAGGCCACTGGGCCCCACGCCCTCCGGCCGCCACCCTGCGC
CGCACCCACAGTTCAACAGCGCGGAGGCCCGGCTGGGGACCGCCACCGCGGCTGCCACGCCCGCCGGGCACA
GACTTGGCCACCTCCTCCCCATGGGGGGGCGGACAGGACTGCGCCCCCGTGCCCTAGGCCGGGGGCCCCCG
ATGCCTTGGCAGTGCCAGCCACGGGAACCAGGAGCGAGAGACGGTGCCAGAACGCCGGGGCCCCGGGCAACTCCG
AGTGGGTGCTCAAGTCCCCCGCGACCCACCGCGGAGTGGGGGGCCCCCTCCGCCACAAGGAAGCACAAACCAG
CTCGCCCTCCCCCTACCCGGGGCGCGAGGACGCTGAGACGGTTTGGGGGTGGGTGGGCGGGAGGACTTTGCTATG
GATTTGAGGTTGACCTTATGCGCGTAGGTTTTGGTTTTTTTTTGCAGTTTTGGTTTTCTTTGCGGTTTTCTAAC
AATTGCACAACTCCGTTCTCGGGGTGGCGGCAGGCGAGGGAGGCTTGGACGCCGTTGGGAATGGGGGGCCACAG
CTGCAGACCTAAGCCCCCCCCACCCCTGGAAAGGTCCCTCCCCAACCCAGGCCCTGGCGTGTGTGGGTGTGCG
TGCGTGTGCGTGCCGTGTTCTGTGTGCAAGGGGCGGGGAGGTGGGCGTGTGTGTGCGTGCCAGCGAAGGCTGCTG
TGGGCGTGTGTGTCAAGTGGGCCACGCGTGCAGGCTGTGTGTCCACGAGCGACGATCGTGGTGGCCCCAGCGGCC
TGGGCGTTGGCTGAGCCGACGCTGGGGCTTCCAGAAGGCCCGGGGTCTCCGAGGTGCCGGTATGAGGTTTGAAC
CCCCCCCCACTCTGCAGAGGGAAGCGGGGACAATGCCGGGTTTTCAGGCAGGAGACACGAGGAGGGCCTGCCCGGA
AGTCACATCGGCAGCAGCTGTCTAAAGGGCTTGGGGCCCTGGGGGGCGGCGAAGGTGGGTGGGGCCCCCTCTGTAA
ATACGGCCCCAGGGTGGTGAGAGATCCCATGCCACCGTCCCCCTTGTGACCTCCCCCTATGACCTCCAGCTGA
CCATGCATGCCACGTGGCTGGCTGGGTCTCTGCCCTCTTTGGAGTTTGCCTCCCCAGCCCCCTCCCCATCAAT
AAAACCTCTGTTTACAACCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 16

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA80145

><subunit 1 of 1, 888 aa, 1 stop

><MW: 95285, pI: 8.89, NX(S/T): 8

MQTPRASPPRPALLLLLLLLGGAHGLFPEEPPPLSVAPRDYLNHYPVFVSGSGPGRLTPAEGA
DDLNIQVRVLRVNRTLFIGDRDNLYRVELEPPTSTELRYQRKLTWRSNPSDINVCRMKGKQEG
ECRNFKVLLLRDESTLFVCGSNAFNPVCANYSIDTLQPVGDNISGMARCPYDPKHANVALF
SDGMLFTATVTDFLAIDAVIYRSLGDRPTLRTVKHDSKWFKEPYFVHAVEWGSFVYFFFREI
AMEFNYLEKVVVSRVARVCKNDVGGSPRVLEKQWTSFLKARLNCSVPGDSHFYFNVLQAVTG
VVSLLGGRPVVLAVFSTPSNSIPGSAVCAFDLTQVAAVFEGRFREQKSPESIWTVPEDQVPR
PRPGCCAAPGMQYNASSALPDDILNFVKTHPLMDEAVPSLGHAPWILRTLMRHQLTRVAVDV
GAGPWGNQTVVFLGSEAGTVLKFLVRPNASTSGTSGLSVFLEEFETYRPDRCGRPGGGETGQ
RLLSLELDAASGGLLAAPRCVVRVPVARCQYSGCMKNCIGSQDPYCGWAPDGSCIFLSPG
TRAAFEQDVSGASTSGLGDCTGLLRASLSEDRAGLVSVNLLVTSSVAAFVVGAVVSGFSVGW
FVGLRERRELARRKDKEAILAHGAGEAVLSVSR LGERRAQGPGRGGGGGGGAGVPPEALLA
PLMQNGWAKATLLQGGPHDLDSGLLPTPEQTPLPQKRLPTPHPHPHALGPRAWDHGHPLLP
SASSSLLLAPARAPEQPPAPGEPTPDGRLYAARPGRASHGDFPLTPHASPDRRRVVSAPTG
PLDPASAADGLPRPWSPPTGSLRRPLGPHAPPAATLRRTHTFNSGEARPGDRHRGCHARPG
TDLAHLHPYGGADRTAPPVP

Important features of the protein:

Signal peptide:

amino acids 1-25

Transmembrane domains:

amino acids 318-339, 598-617

N-glycosylation sites.

amino acids 74-78, 155-159, 167-171, 291-295, 386-390, 441-445,
462-466

Glycosaminoglycan attachment sites.

amino acids 51-55, 573-577

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 102-106

N-myristoylation sites.

amino acids 21-27, 50-56, 189-195, 333-339, 382-388, 448-454,
490-496, 491-497, 508-514, 509-515, 531-537, 558-564, 569-575,
574-580, 580-586, 610-616, 643-649, 663-669, 666-672, 667-673,
668-674, 669-675, 670-676, 868-874, 879-885

FIGURE 17

AGCAACTCAAGTTCATCATTGTCCTGAGAGAGAGGAGCAGCGCGGTCTCGGCCGGGACAGC
AGAACGCCAGGGGACCCTCACCTGGGCGCGCCGGGGCACGGGCTTTGATTGTCTGGGGTCG
CGGAGACCCGCGCGCCTGCCCTGCACGCCGGGCGGGCAACCTTTGCAGTCGCGTTGGCTGCTG
CGATCGGCCGGCGGGTCCCTGCCGAAGGCTCGGCTGCTTCTGTCCACCTCTTACACTTCTTC
ATTTATCGGTGGATCATTTTCGAGAGTCCGTCTTGTAATGTTTGGCACTTTGCTACTTTATT
GCTTCTTTCTGGCGACAGTTCAGCACTCGCCGAGACCGGCGGAGAAAGGCAGCTGAGCCCG
GAGAAGAGCGAAATATGGGGACCCGGGCTAAAAGCAGACGTCGTCTTCCCGCCCGCTATTT
CTATATTCAGGCAGTGGATACATCAGGGAATAAATTCACATCTTCTCCAGGCGAAAAGGTCT
TCCAGGTGAAAGTCTCAGCACCAGAGGAGCAATTCCTAGAGTTGGAGTCCAGGTTTTAGAC
CGAAAAGATGGGTCTTCATAGTAAGATACAGAATGTATGCAAGCTACAAAAATCTGAAGGT
GGAAATTAAATTCCAAGGGCAACATGTGGCCAAATCCCCATATATTTTAAAAGGGCCGGTTT
ACCATGAGAACTGTGACTGTCTCTGCAAGATAGTGCAGCCTGGCTACGGGAGATGAACTGC
CCTGAAACCATTTGCTCAGATTCAGAGAGATCTGGCACATTTCCCTGCTGTGGATCCAGAAAA
GATTGCAGTAGAAATCCCAAAAAGATTTGGACAGAGGCAGAGCCTATGTCACTACACCTTAA
AGGATAACAAGGTTTATATCAAGACTCATGGTGAACATGTAGGTTTTAGAATTTTCATGGAT
GCCATACTACTTTCTTTGACTAGAAAGGTGAAGATGCCAGATGTGGAGCTCTTTGTTAATTT
GGGAGACTGGCCTTTGGAAAAAAGAAATCCAATTCAAACATCCATCCGATCTTTTCCTGGT
GTGGCTCCACAGATTCCAAGGATATCGTGATGCCTACGTACGATTTGACTGATTCTGTTCTG
GAAACCATGGGCCGGGTAAGTCTGGATATGATGTCCGTGCAAGCTAACACGGGTCCTCCCTG
GGAAAGCAAAAATTCCTACTGCCGTCTGGAGAGGGCGAGACAGCCGCAAAGAGAGACTCGAGC
TGTTTAACTCAGTAGAAAACACCCAGAACTCATAGACGCTGCTTTCACCAACTTTTTCTTC
TTTAAACACGATGAAAACCTGTATGGTCCCATTGTGAAACATATTTTCAATTTTTTGATTTCTT
CAAGCATAAGTATCAAATAAATATCGATGGCACTGTAGCAGCTTATCGCCTGCCATATTTGC
TAGTTGGTGACAGTGTTGTGCTGAAGCAGGATTCCATCTACTATGAACATTTTTTACAATGAG
CTGCAGCCCTGGAAACACTACATTCCAGTTAAGAGCAACCTGAGCGATCTGCTAGAAAACT
TAAATGGGCGAAAGATCACGATGAAGAGGCCAAAAAGATAGCAAAAGCAGGACAAGAATTTG
CAAGAAATAATCTCATGGGCGATGACATATTCTGTTATTATTTCAAACCTTTTCCAGGAATAT
GCCAATTTACAAGTGAGTGAGCCCCAAATCCGAGAGGGCATGAAAAGGGTAGAACCACAGAC
TGAGGACGACCTCTTCCCTTGTAATTGCCATAGGAAAAAGACCAAAGATGAACTCTGATATG
CAAAATAACTTCTATTAGAATAATGGTGCTCTGAAGACTCTTCTTAACTAAAAAGAAGAATT
TTTTTAAGTATTAATTCATGGACAATATAAAATCTGTGTGATTGTTTGCAGTATGAAGACA
CATTTCTACTTATGCAGTATTCTCATGACTGTACTTTAAAGTACATTTTTTAGAATTTTATAA
TAAAACCACCTTTATTTTAAAGGAAAAAA

FIGURE 18

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA84917

><subunit 1 of 1, 502 aa, 1 stop

><MW: 58043, pI: 7.94, NX(S/T): 2

MFGTLLLYCFFLATVPALAETGGERQLSPEKSEIWGPGLKADVVLPAFYFYIQAVDTSGNKF
TSSPGEKVFQVKVSAPEEQFTRVGVQVLDRKDGSFIVRYRMYASYKNLKVVEIKFQGQHVAKS
PYILKGPVYHENCDCPLQDSAAWLREMNCPETIAQIQRDLAHFPAVDPEKIAVEIPKRFGQR
QSLCHYTLKDNKVYIKTHGEHVGFRIFMDAILLSLTRKVKMPDVELFVNLGDWPLEKKKSNS
NIHPIFSWCGSTDSDKDIVMPTYDLTDSVLETMGRVSLDMMSVQANTGPPWESKNSTAVWRGR
DSRKERLELVKLSRKHPOLIDAAFTNFFFFKHDENLYGPIVKHISFFDFFKHKYQINIDGTV
AAYRLPYLLVGDSVVLKQDSIYYEHFYNELQPWKHYIPVKSNLSDLLEKLKWAKDHDEEAKK
IAKAGQEFARNNLMGDDIFCYFVKLFQEYANLQVSEPQIREGMKRVEPQTEDDLFPCTCHRK
KTKDEL

Important features of the protein:

Signal peptide:

amino acids 1-17

N-glycosylation sites.

amino acids 302-306, 414-418

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 243-247, 495-499

Tyrosine kinase phosphorylation site.

amino acids 341-348

N-myristoylation sites.

amino acids 59-65, 118-124, 184-190, 258-264, 370-376, 439-445

Endoplasmic reticulum targeting sequence.

amino acids 499-504

FIGURE 19

CCTGGAGCCGGAAGCGCGGCTGCAGCAGGGCGAGGCTCCAGGTGGGGTTCGGTTCCGCATCCA
GCCTAGCGTGTCCACG**ATG**CGGCTGGGCTCCGGGACTTTCGCTACCTGTTGCGTAGCGATCG
AGGTGCTAGGGATCGCGGTCTTCCTTCGGGGATTCTTCCCGGCTCCCGTTCGTTCTCTGCC
AGAGCGGAACACGGAGCGGAGCCCCAGCGCCCGAACCTCGGCTGGAGCCAGTTCTAACTG
GACCACGCTGCCACCACCTCTCTTCAGTAAAGTTGTTATTGTTCTGATAGATGCCTTGAGAG
ATGATTTTGTGTTTGGGTCAAAGGGTGTGAAATTTATGCCCTACACAACCTACCTTGTTGGAA
AAAGGAGCATCTCACAGTTTGTGGCTGAAGCAAAGCCACCTACAGTTACTATGCCTCGAAT
CAAGGCATTGATGACGGGGAGCCTTCCTGGCTTTGTGACGTCATCAGGAACCTCAATTCTC
CTGCACTGCTGGAAGACAGTGTGATAAGACAAGCAAAAGCAGCTGGAAAAAGAATAGTCTTT
TATGGAGATGAAACCTGGGTAAATTTATCCCAAAGCATTTTGTGGAATATGATGGAACAAC
CTCATTTTTCTGTGTCAGATTACACAGAGGTGGATAATAATGTCACGAGGCATTTGGATAAAG
TATTAAAAAGAGGAGATTGGGACATATTAATCCTCCACTACCTGGGGCTGGACCACATTGGC
CACATTTTCAGGGCCCAACAGCCCCCTGATTGGGCAGAAGCTGAGCGAGATGGACAGCGTGCT
GATGAAGATCCACACCTCACTGCAGTCGAAGGAGAGAGAGACGCCTTTACCCAATTTGCTGG
TTCTTTGTGGTGACCATGGCATGTCTGAAACAGGAAGTCACGGGGCCTCCTCCACCGAGGAG
GTGAATACACCTCTGATTTTAATCAGTTCTGCGTTTGAAAGGAAACCCGGTGATATCCGACA
TCCAAAGCACGTCCA**ATAG**ACGGATGTGGCTGCGACACTGGCGATAGCACTTGGCTTACCGA
TTCCAAAAGACAGTGTAGGGAGCCTCCTATTCCAGTTGTGGAAGGAAGACCAATGAGAGAG
CAGTTGAGATTTTACATTTGAATACAGTCAGCTTAGTAAACTGTTGCAAGAGAATGTGCC
GTCATATGAAAAAGATCCTGGGTTTGAGCAGTTTAAATGTGAGAAAGATTGCATGGGAAC
GGATCAGACTGTACTTGGAGGAAAAGCATTGAGAAGTCCTATTCAACCTGGGCTCCAAGGTT
CTCAGGCAGTACCTGGATGCTCTGAAGACGCTGAGCTTGTCCTGAGTGACAAAGTGCCCCA
GTTCTCACCCCTGCTCCTGCTCAGCGTCCCACAGGCACTGCACAGAAAGGCTGAGCTGGAAGT
CCCCTGTCTCTCCTGGGTTTTCTCTGCTCTTTTATTGTTGGTGATCCTGGTTCTTTTCGGCCG
TTCACGTCAATTGTGTGCACCTCAGCTGAAAGTTCGTGCTACTTCTGTGGCCTCTCGTGGCTG
GCGGCAGGCTGCCTTTCTGTTTACCAGACTCTGGTTGAACACCTGGTGTGTGCCAAGTGCTGG
CAGTGCCCTGGACAGGGGGCCTCAGGGAAGGACGTGGAGCAGCCTTATCCAGGCCTCTGGG
TGTCCCGACACAGGTGTTTACATCTGTGCTGTGAGGTCAGATGCCTCAGTTCTTGGAAGCT
AGGTTCTGCGACTGTTACCAAGGTGATTGTAAAGAGCTGGCGGTCACAGAGGAACAAGCCC
CCCAGCTGAGGGGGTGTGTGAATCGGACAGCCTCCCAGCAGAGGTGTGGGAGCTGCAGCTGA
GGGAAGAAGAGACAATCGGCCTGGACACTCAGGAGGGTCAAAGGAGACTTGGTTCGCACCAC
TCATCCTGCCACCCCCAGAATGCATCCTGCCTCATCAGGTCCAGATTTCTTTCCAAGGCGGA
CGTTTTCTGTTGGAATTCTTAGTCCTTGGCCTCGGACACCTTCATTTCGTTAGCTGGGGAGTG
GTGGTGAGGCAGTGAAGAAGAGGCGGATGGTCACACTCAGATCCACAGAGCCCAGGATCAAG
GGACCCACTGCAGTGGCAGCAGGACTGTTGGGCCCCCACCCTGCACAGCCCTCATC
CCCTCTTGGCTTGAGCCGTGAGAGGCCCTGTGCTGAGTGTCTGACCGAGACACTCACAGCTT
TGTCATCAGGGCACAGGCTTCCTCGGAGCCAGGATGATCTGTGCCACGCTTGACCTCGGGC
CCATCTGGGCTCATGCTCTCTCTCCTGCTATTGAATTAGTACCTAGCTGCACACAGTATGTA
GTTACCAAAGAATAAACGGCAATAATTGAGAAAAAAA

FIGURE 20

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA84920

><subunit 1 of 1, 310 aa, 1 stop

><MW: 33875, pI: 7.08, NX(S/T): 2

MRLGSGTFATCCVAIEVLGIAVFLRGFFPAPVRSSARAEGAEPPAPEPSAGASSNWTTLPP
PLFSKVVIVLIDALRDDFVFGSKGVKFMPYTTYLVEKGASHSFVAEAKPPTVTMPRIKALMT
GSLPGFVDVIRNLNSPALLEDVIRQAKAAGKRIVFYGDETWVKLFPKHFVEYDGTTSFFVS
DYTEVDNNVTRHLDKVLKRGDWDILILHYLGLDGHIGHISGPN SPLIGQKLSEMDSVLMKIHT
SLQSKERETPLPNLLVLCGDHGMSETGSHGASSTEEVNTPLILISSAFERKPGDIRHPKHVQ

Important features of the protein:

Signal peptide:

amino acids 1-34

Transmembrane domain:

amino acids 58-76

N-glycosylation sites.

amino acids 56-60, 194-198

N-myristoylation sites.

amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,
275-281, 278-284

Amidation site.

amino acids 154-158

Cell attachment sequence.

amino acids 205-208

FIGURE 21

AGCCAGGCAGCACATCACAGCGGGAGGAGCTGTCCCAGGTGGCCCAGCTCAGCA**ATGG**CAAT
GGGGGTCCCCAGAGTCATTCTGCTCTGCCTCTTTGGGGCTGCGCTCTGCCTGACAGGGTCCC
AAGCCCTGCAGTGCTACAGCTTTGAGCACACCTACTTTGGCCCCTTTGACCTCAGGGCCATG
AAGCTGCCCAGCATCTCCTGTCCTCATGAGTGCTTTGAGGCTATCCTGTCTCTGGACACCGG
GTATCGCGCGCCGGTGACCCTGGTGCGGAAGGGCTGCTGGACCGGGCCTCCTGCGGGCCAGA
CGCAATCGAACCCGGACGCGCTGCCGCCAGACTACTCGGTGGTGCGCGGCTGCACAACCTGAC
AAATGCAACGCCCACCTCATGACTCATGACGCCCTCCCCAACCTGAGCCAAGCACCCGACCC
GCCGACGCTCAGCGGCGCCGAGTGCTACGCCTGTATCGGGGTCCACCAGGATGACTGCGCTA
TCGGCAGGTCCCGACGAGTCCAGTGTCACCAGGACCAGACCGCCTGCTTCCAGGGCAGTGGC
AGAATGACAGTTGGCAATTTCTCAGTCCCTGTGTACATCAGAACCTGCCACCGGCCCTCCTG
CACCACCGAGGGCACCACCAGCCCCCTGGACAGCCATCGACCTCCAGGGCTCCTGCTGTGAGG
GGTACCTCTGCAACAGGAAATCCATGACCCAGCCCTTCACCAGTGCTTCAGCCACCACCCCT
CCCCGAGCACTACAGGTCCTGGCCCTGCTCCTCCCAGTCCTCCTGCTGGTGGGGCTCTCAGC
ATAGACCGCCCCCTCCAGGATGCTGGGGACAGGGCTCACACACCTCATTCTTGCTGCTTCAGC
CCCTATCACATAGCTCACTGGAAAATGATGTTAAAGTAAGAATTGCAAAA

FIGURE 22

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA86576

><subunit 1 of 1, 251 aa, 1 stop

><MW: 26935, pI: 7.42, NX(S/T): 2

MAMGVPRVILLCLFGAALCLTGSQALQCYSFEHTYFGPFDLRAMKLPSISCPHECFEAILSL
DTGYRAPVTLVRKGCWTGPPAGQTQSNPDALPPDYSVVRGCTTDKCNAHLMTHDALPNLSQA
PDPPTLSGAECYACIGVHQDDCAIGRSRRVQCHQDQTACFQGSGRMTVGNFSPVYIIRTCHR
PSCTTEGTTSPWTAIDLQGSCEGYLCNRKSMTQPFTSASATTPPRALQVLALLLPVLLLVGLSA

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 233-251

N-glycosylation sites.

amino acids 120-124, 174-178

N-myristoylation sites.

amino acids 15-21, 84-90

FIGURE 23

CCCACGCGTCCGGGACAGATGAACTTAAAAGAGAAGCTTTAGCTGCCAAAGATTGGGAAAGG
GAAAGGACAAAAAAGACCCCTGGGCTACACGGCGTAGGTGCAGGGTTTCCTACTGCTGTTCT
TTTATGCTGGGAGCTGTGGCTGTAACCAACTAGGAAATAACGTATGCAGCAGCTATGGCTGT
CAGAGAGTTGTGCTTCCCAAGACAAAGGCAAGTCCTGTTTCTTTTTCTTTTTTGGGGAGTGT
CCTTGGCAGGTTCTGGGTTTGGACGTTATTCGGTGACTGAGGAAACAGAGAAAGGATCCTTT
GTGGTCAATCTGGCAAAGGATCTGGGACTAGCAGAGGGGGAGCTGGCTGCAAGGGGAACCAG
GGTGGTTTCCGATGATAACAAACAATACCTGCTCCTGGATTACATACCGGGAATTTGCTCA
CAAATGAGAACTGGACCGAGAGAAGCTGTGTGGCCCTAAAGAGCCCTGTATGCTGTATTTT
CAAATTTTAATGGATGATCCCTTTCAGATTTACCGGGCTGAGCTGAGAGTCAGGGATATAAA
TGATCACGCGCCAGTATTTTCCAGGACAAAGAAACAGTCTTAAAAATATCAGAAAATACAGCTG
AAGGGACAGCATTTAGACTAGAAAGAGCACAGGATCCAGATGGAGGACTTAACGGTATCCAA
AACTACACGATCAGCCCCAAGTCTTTTTTCCATATTAACATTAGTGGCGGTGATGAAGGCAT
GATATATCCAGAGCTAGTGTGGACAAAGCACTGGATCGGGAGGAGCAGGGAGAGCTCAGCT
TAACCCTCACAGCGCTGGATGGTGGGTCTCCATCCAGGTCTGGGACCTCTACTGTACGCATC
GTTGTCTTGGACGTCATGACAATGCCCCACAGTTTGCCAGGCTCTGTATGAGACCCAGGC
TCCAGAAAACAGCCCCATTGGGTTCTTATTGTTAAGGTATGGGCAGAAGATGTAGACTCTG
GAGTCAACGCGGAAGTATCCTATTCATTTTTTGTATGCCTCAGAAAATATTCGAACGACCTTT
CAAATCAATCCTTTTTCTGGGGAAATCTTTCTCAGAGAATTGCTTGATTATGAGTTAGTAAA
TTCTTACAAAATAAATATACAGGCAATGGACGGTGGAGGCCTTTCTGCAAGATGTAGGGTTT
TAGTGGAAGTATTGGACACCAATGACAATCCCCCTGAACTGATCGTATCATCATTTTTCCAAC
TCTGTTGCTGAGAATTCTCCTGAGACGCCGCTGGCTGTTTTTAAGATTAATGACAGAGACTC
TGGAGAAAATGGAAAGATGTTTTGCTACATTCAAGAGAATCTGCCATTCTACTAAAACCTT
CTGTGGAGAATTTTTACATCCTAATTACAGAAGGCGCGCTGGACAGAGAGATCAGAGCCGAG
TACAACATCACTATCACCGTCACTGACTTGGGGACACCCAGGCTGAAAACCGAGCACAAACAT
AACGGTCTTGGTCTCCGACGTCATGACAACGCCCGCCCTTACCCAAACCTCCTACACCC
TGTTTCGTCGCGAGAACAAACAGCCCCGCCCTGCACATCGGCAGCGTCAGCGCCACAGACAGA
GACTCGGGACCAACGCCAGGTCACTACTCGCTGCTGCCGCCCAAGACCCGCACCTGCC
CCTCGCCTCCCTGGTCTCCATCAACGCGGACAACGGCCACCTGTTTCGCCCTCAGGTGCTGG
ACTACGAGGCCCTGCAGGCTTTCGAGTTCCGCGTGGGCGCCACAGACCGCGGCTCCCCCGCG
CTGAGCAGAGAGGCGCTGGTGGCGGTGCTGGTGGTGGACGCCAACGACAACCTCGCCCTTCGT
GCTGTACCCGCTGCAGAACGGCTCCGCGCCCTGCACCGAGCTGGTGCCCCGGGCGGCGGAGC
CGGGCTACCTGGTGACCAAGGTGGTGGCGGTGGACGGCGACTCGGGCCAGAACGCCTGGCTG
TCGTACCAGCTGCTCAAGGCCACGGAGCCCGGGCTGTTTCGGTGTGTGGGCGCACAAATGGGGA
GGTGCGCACCGCCAGGCTGCTGAGCGAGCGCGACGACGCAAGCACAGGCTCGTGGTGTGTTG
TCAAGGACAATGGCGAGCCTCCTCGCTCGGCCACCGCCACGCTGCACTTGCTCCTGGTGGAC
GGCTTCTCCAGCCCTACCTGCCTCTCCCGAGGCGGCCCGGCCAGGCCCAGGCCCAGGCGAGGC
CGACTTGCTCACCGTCTACCTGGTGGTGGCGTTGGCCTCGGTGTCTTCGCTCTTCCTCCTCT
CGGTGCTCCTGTTTCGTGGCGGTGCGGCTGTGCAGGAGGAGCAGGGCGGCCTCGGTGGGTGCG
TGCTCGGTGCCCGAGGGTCTTTTCCAGGGCATCTGGTGGACGTGAGGGGCGCTGAGACCCT
GTCCAGAGCTACCAGTATGAGGTGTGTCTGACGGGAGGCCCGGGACCAGTGAGTTCAAGT
TCTTGAAACCAGTTATTTTCGGATATTCAGGCACAGGGCCCTGGGAGGAAGGTGAAGAAAAT
TCCACCTTCCGAAATAGCTTTGGATTTAATATTCAGTAAAGTCTGTTTTTAGTTTCATATAC
TTTTGGTGTGTACATAGCCATGTTTCTATTAGTTTACTTTTAAATCTCAAATTTAAGTTAT
TATGCAACTTCAAGCATTATTTTCAAGTAGTATACCCCTGTGGTTTTTACAATGTTTCATCAT
TTTTTTGCATTAATAACAACCTGGGTTTAAATTTAATGAGTATTTTTTTCTAAATGATAGTGT
AAGGTTTTAATTCCTTCCAAGTGGCAAGGAATTAATTACTATTATATCTCATTACAGAAAT
CTGAGGTTTTGATTCATTTTCAAGCTTGCATCTCATGATTCTAATCACTTCTGTCTATAGTG
TACTTGCTCTATTTAAGAAGGCATATCTACATTTCCAAACTCATTCTAACATTCTATATATT
CGTGTTTGAAAACCATGTCAATTTATTTCTACATCATGTATTTAAAAAGAAATATTTCTCTAC
TACTATGCTCATGACAAAATGAAACAAAGCATAATTGTGAGCAATACTGAACATCAATAATAC
CCTTAGTTTATATACTTATTATTTTCTTTAAGCATGCTACTTTTACTTGGCCAATATTTT
CTTATGTTAACTTTTGTGATGTATAAAACAGACTATGCCTTATAATTGAAATAAAATTTATA
ATCTGCCTGAAAATGAATAAAAATAAAACATTTTGAAATGTGAAAAAAAAAAAAAAAAAAAAA

FIGURE 24

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA87976

><subunit 1 of 1, 800 aa, 1 stop

><MW: 87621, pI: 4.77, NX(S/T): 7

MAVRELCFPRQRQVLFLFLFWGVSLAGSGFGGRYSVTEETEKGSFVVNLAKDLGLAEGELAAR
GTRVVSDDNKQYLLLDSTGNLLTNEKLDREKLCGPKEPCMLYFQILMDDPFQIYRAELRVR
DINDHAPVFQDKETVLKISENTAEGTAFRLERAQDPDGGGLNGIQNYTISPNSFFHINISGGD
EGMIYPELVLDKALDREEQGELSLLTALDGGSPSRSGTSTVRIVVLDVNDNAPQFAQALYE
TQAPENSPIGFLIVKVWAEDVDSGVNAEVSYSFFDASENIRTTTFQINPFSGEIFLRELLDYE
LVNSYKINIQAMDGGGLSARCRVLVEVLDTNDNPPELIVSSFSNSVAENSPETPLAVFKIND
RDSGENGKMVCYIQENLPFLLKPSVENFYILITEGALDREIRAENITITVTDLGTPLRKTE
HNITVLVSDVNDNAPAFQTQTSYTLFVRENNSPALHIGSVSATDRDSTNAQVTYSLLPPQDP
HLPLASLV SINADNGHLFALRSLDYEALQAFEFVRGATDRGSPALSREALVRVLVLDANDNS
PFVLYPLQNGSAPCTELVPRAAEPGYLVTKVAVDGDGSGQNAWLSYQLLKATEPGLFGVWAH
NGEVRTARLLSERDAAKHRLVVLVKDNGEPPRSATATLHLLLVDFGFSQPYLPLPEAAPAQAO
AEADLLTVYLVVALASVSSLFLLSVLLFVAVRLCRRSRAASVGRCSVPEGPFPGHLVDVRGA
ETLSQSYQYEVCLTGGPGTSEFKFLKPVISDIQAQGPGRKGEENSTFRNSFGFNIQ

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 687-711

N-glycosylation sites.

amino acids 169-173, 181-185, 418-422, 436-440, 567-571, 788-792

Glycosaminoglycan attachment site.

amino acids 28-32

Tyrosine kinase phosphorylation sites.

amino acids 394-402, 578-585

N-myristoylation sites.

amino acids 22-28, 27-33, 53-59, 82-88, 162-168, 184-190,
217-223, 324-330, 325-331, 471-477, 568-574, 759-765

Amidation site.

amino acids 781-785

Aminoacyl-transfer RNA synthetases class-II signature 1.

amino acids 117-138

Cadherins extracellular repeated domain signature.

amino acids 121-132, 230-241, 335-346, 439-450, 549-560

[illegible][illegible]

FIGURE 26

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA92234

><subunit 1 of 1, 507 aa, 1 stop

><MW: 56692, pI: 5.22, NX(S/T): 3

MDPKLGRMAASLLAVLLLLLLERGMFSSPSPPPALLEKVVFQYIDLHQDEFVQTLKEWVAIESD
SVQPVPRFRQELFRMMAVAADTLQRLGARVASVDMGPQQLPDQGQSLPIPPVILAELGSDPTK
GTVCFYGHLDVQPADRGDGWLTDPYVLTEVDGKLYGRGATDNKGPVLAWINAVSAFRALEQD
LPVNIKFIIIEGMEEAGSVALEELVEKEKDRFFSGVDYIVISDNLWISQRKPAITYGTRGNSY
FMVEVKCRDQDFHSGTFGGILHEPMADLVALLGSLVDSSGHILVPGIYDEVVPLTEEEINTY
KAIHLDLEEYRNSSRVEKFLFDTKEEILMHLWRYPSSLIHGIEGAFDEPGTKTVIPGRVIGK
FSIRLVPHMNVSAVEKQVTRHLEDVFSKRNSSNKMVVSMTLGLHPWIANIDDTQYLAAKRAI
RTVFGTEPDMIRDGSTIPIAKMFQÉIVHKSVVLIPLGAVDDGEHSQNEKINRWNYIEGTKLF
AAFFLEMAQLH

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

273-292

N-glycosylation sites.

amino acids 322-326, 382-386, 402-406

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 400-404

N-myristoylation sites.

amino acids 89-95, 119-125, 162-168, 197-203, 242-248, 263-269,
351-357

Cell attachment sequence.

amino acids 140-143

ArgE / dapE / ACY1 / CPG:

amino acids 156-167

FIGURE 27

CTCGGCTGGATTTAAGGTTGCCGCTAGCCGCTGGGAATTTAAGGGACCCACACTACCTTCC
CGAAGTTGAAGGCAAGCGGTGATTGTTTGTAGACGGCGCTTTGTCT**ATG**GGACCTGTGCGGTT
GGGAATATTGCTTTTCCTTTTTTTGGCCGTGCACGAGGCTTGGGCTGGGATGTTGAAGGAGG
AGGACGATGACACAGAACGCTTGCCAGCAAATGCGAAGTGTGTAAGCTGCTGAGCACAGAG
CTACAGGCGGAAGTGAAGTGCACCGGTGATCTCGAGAGGTGCTGGAGCTGGGGCAGGTGCT
GGATACAGGCAAGAGGAAGAGACACGTGCCTTACAGCGTTTCAGAGACAAGGCTGGAAGAGG
CCTTAGAGAATTTATGTGAGCGGATCCTGGACTATAGTGTTACAGCTGAGCGCAAGGGCTCA
CTGAGATATGCCAAGGGTCAGAGTCAGACCATGGCAACACTGAAAGGCCTAGTGCAGAAGGG
GGTGAAGGTGGATCTGGGGATCCCTCTGGAGCTTTGGGATGAGCCCAGCGTGGAGGTCACAT
ACCTCAAGAAGCAGTGTGAGACCATGTTGGAGGAGTTTGAAGACATTGTGGGAGACTGGTAC
TTCCACCATCAGGAGCAGCCCCACAAAATTTTCTCTGTGAAGGTCATGTGCTCCCAGCTGC
TGAAACTGCATGTCTACAGGAACTTGGACTGGAAAGGAGATCACAGATGGGGAAGAGAAAA
CAGAAGGGGAGGAAGAGCAGGAGGAGGAGGAGGAAGAGGAGGAAGAGGAAGGGGGAGACAAG
ATGACCAAGACAGGAAGCCACCCCAAAGTTGACCGAGAAGATCTTT**TGA**CCCTTGCCTTTGAG
CCCCCAGGAGGGGAAGGGATCATGGAGAGCCCTCTAAAGCCTGCACTCTCCCTGCTCCACAG
CTTTCAGGGTGTGTTTATGAGTGACTCCACCCAAGCTTGTAGCTGTTCTCTCCCATCTAACC
TCAGGCAAGATCCTGGTGAACAGCATGACATGGCTTCTGGGGTGGAGGGTGGGGGTGGAGG
TCCTGCTCCTAGAGATGAACTCTATCCAGCCCCCTAATTGGCAGGTGTATGTGCTGACAGTA
CTGAAAGCTTTCTCTTTAACTGATCCCACCCCCACCCAAAAGTCAGCAGTGGCACTGGAGC
TGTGGGCTTTGGGGAAGTCACTTAGCTCCTTAAGGTCTGTTTTTAGACCCTTCCAAGGAAGA
GGCCAGAACGGACATTCTCTGCGATCTATATACATTGCCTGTATCCAGGAGGCTACACACCA
GCAAACCGTGAAGGAGAATGGGACACTGGGTGATGGCCTGGAGTTGCTGATAATTTAGGTGG
GATAGATACTTGGTCTACTTAAGCTCAATGTAACCCAGAGCCCACCATATAGTTTTATAGGT
GCTCAACTTCTATATCGCTATTAACTTTTTTCTTTTTTTCTA

FIGURE 28

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA92256

><subunit 1 of 1, 248 aa, 1 stop

><MW: 28310, pI: 4.63, NX(S/T): 0

MGPVRLGILLFLFLAVHEAWAGMLKEEDDDTERLPSKCEVCKLLSTELQAELSRTGRSREVL
ELGQVLDTGKRKRHVPYSVSETRLEEALENLCERILDYSVHAERKGSRLRYAKGQSQTMATLK
GLVQKGVKVDLGIPLLEWDEPSVEVTYLKKQCETMLEEFEDIVGDWYFHHQEQLQNFLCEG
HVLPAAETACLQETWTGKEITDGEEKTEGEEEEEEEEEEEEEGGDKMTKTGSHPKLDREDL

Important features of the protein:

Signal peptide:

amino acids 1-21

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 106-110

N-myristoylation site.

amino acids 115-121

Amidation site.

amino acids 70-74

FIGURE 29

AAGTACTTGTGTCCGGGTGGTGGACTGGATTAGCTGCGGAGCCCTGGAAGCTGCCTGTCCTT
CTCCCTGTGCTTAACCAGAGGTGCCCATGGGTTGGACAATGAGGCTGGTCACAGCAGCACTG
TTACTGGGTCTCATGATGGTGGTCACTGGAGACGAGGATGAGAACAGCCCCTGTGCCCATGA
GGCCCTCTTGGACGAGGACACCCTCTTTTGCCAGGGCCTTGAAGTTTTCTACCCAGAGTTGG
GGAACATTGGCTGCAAGGTTGTTCCCTGATTGTAACAACCTACAGACAGAAGATCACCTCCTGG
ATGGAGCCGATAGTCAAGTTCCCGGGGGCCGTGGACGGCGCAACCTATATCCTGGTGATGGT
GGATCCAGATGCCCCCTAGCAGAGCAGAACCCAGACAGAGATTCTGGAGACATTGGCTGGTAA
CAGATATCAAGGGCGCCGACCTGAAGAAAGGGAAGATTCAGGGCCAGGAGTTATCAGCCTAC
CAGGCTCCCTCCCCACCGGCACACAGTGGCTTCCATCGCTACCAGTTCTTTGTCTATCTTCA
GGAAGGAAAAGTCATCTCTCTCCTTCCCAAGGAAAACAAAACCTCGAGGCTCTTGGAATGG
ACAGATTTCTGAACCGCTTCCACCTGGGCGAACCTGAAGCAAGCACCAGTTCATGACCCAG
AACTACCAGGACTCACCAACCCTCCAGGCTCCAGAGGAAGGGCCAGCGAGCCCAAGCACAA
AACCAGGCAGAGATTAGCTGCCTGCTAGATAGCCGGCTTTGCCATCCGGGCATGTGGCCACAC
TGCTCACCACCGACGATGTGGGTATGGAACCCCTCTGGATACAGAACCCCTTCTTTTCCAA
ATTAAAAAAAAAATCATCAA

FIGURE 30

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA92274

><subunit 1 of 1, 223 aa, 1 stop

><MW: 25402, pI: 8.14, NX(S/T): 1

MGWTMRLVTAALLLGLMMVVTGDEDENSPCAHEALLDEDTLFCQGLEVFYPELGNIGCKVVP
DCNNYRQKITSWMEPIVKFPGAVDGATYILVMVDPDAPSRAEPRQRFWRHWLVTDIKGADLK
KGKIQGQELSA YQAPSPPAHSGFHYQFFVYLQEGKVISLLPKENKTRGSWKMDRFLNRFHL
GEPEASTQFMTQNYQDSPTLQAPRGRASEPKHKTRQR

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 169-173

Tyrosine kinase phosphorylation site.

amino acids 59-68

N-myristoylation sites.

amino acids 54-60, 83-89, 130-136

Phosphatidylethanolamine signature.

amino acids 113-157

FIGURE 31

GTCGACCCACGCGTCCGAAGCTGCTGGAGCCACGATTTCAGTCCCCTGGACTGTAGATAAAGA
CCCTTTCTTGCCAGGTGCTGAGACAACCACACTATGGAGAGGCACTCCAGGAGACGCTGATGG
TGGAGGAAGGGCCGTCTATCAATCAATCACTGTTGCTGTTATCACATGCAAGTATCCAGAGG
CTCTTGAGCAAGGCAGAGGGGATCCCATTTATTTGGGAATCCAGAATCCAGAAATGTGTTTG
TATTGTGAGAAGGTTGGAGAACAGCCCACATTGCAGCTAAAAGAGCAGAAGATCATGGATCT
GTATGGCCAACCCGAGCCCGTGAAACCCTTCCTTTTCTACCGTGCCAAGACTGGTAGGACCT
CCACCCTTGAGTCTGTGGCCTTCCCGGACTGGTTCATTGCCTCCTCCAAGAGAGACCAGCCC
ATCATTCTGACTTCAGAACTTGGGAAGTCATACAACACTGCCTTTGAATTAAATATAAATGA
CTGAACTCAGCCTAGAGGTGGCAGCTTGGTCTTTGTCTTAAAGTTTCTGGTTCCCAATGTGT
TTTCGTCTACATTTTCTTAGTGTCATTTTCACGCTGGTGCTGAGACAGGAGCAAGGCTGCTG
TTATCATCTCATTTTATAATGAAGAAGAAGCAATTACTTCATAGCAACTGAAGAACAGGATG
TGGCCTCAGAAGCAGGAGAGCTGGGTGGTATAAGGCTGTCCTCTCAAGCTGGTGCTGTGTAG
GCCACAAGGCATCTGCATGAGTGACTTTAAGACTCAAAGACCAAACACTGAGCTTTCTTCTA
GGGGTGGGTATGAAGATGCTTCAGAGCTCATGCGCGTTACCCACGATGGCATGACTAGCACA
GAGCTGATCTCTGTTTCTGTTTTGCTTTATTCCCTCTTGGGATGATATCATCCAGTCTTTAT
ATGTTGCCAATATAACCTCATTGTGTGTAATAGAACCTTCTTAGCATTAAGACCTTGTAACA
AAAATAATTCTTGGGGTGGGTATGAAGATGCTTCAGAGCTCATGCGCGTTACCCACGATGGC
ATGACTAGCACAGAGCTGATCTCTGTTTCTGTTTTGCTTTATTCCCTCTTGGGATGATATCA
TCCAGTCTTTATATGTTGCCAATATAACCTCATTGTGTGTAATAGAACCTTCTTAGCATTAAG
ACCTTGTAACAATAAATAATTCTTGTGTTAAGTTAAATCATTTTTGTCTAATTGTAATGTG
TAATCTTAAAGTTAAATAAACTTTGTGTATTTATATAATAATAAAGCTAAAACCTGATATAAA
ATAAAGAAAGAGTAAACTG

1000
900
800
700
600
500
400
300
200
100
0

FIGURE 32

MRGTPGDADGGGRAVYQSITVAVITCKYPEALEQGRGDPIYLGIONPEMCLYCEKVGEQPTL
QLKEQKIMDLYGQPEPVKPFIFYRAKTGRTSTLESVAFPDWFIASSKRDQPIILTSELGKSY
NTAFELNIND

Signal sequence:

amino acids 1-17

N-myristoylation site.

amino acids 10-16

Cell attachment sequence.

amino acids 36-39

MRGTPGDADGGGRAVYQSITVAVITCKYPEALEQGRGDPIYLGIONPEMCLYCEKVGEQPTL
QLKEQKIMDLYGQPEPVKPFIFYRAKTGRTSTLESVAFPDWFIASSKRDQPIILTSELGKSY
NTAFELNIND

FIGURE 33

GCGAGGCTGCACCAGCGCCTGGCACC**ATG**AGGACGCCTGGGCCTCTGCCCCTGCTGCTGCTG
CTCCTGGCGGGAGCCCCCGCCGCGCGGCCCACTCCCCGACCTGCTACTCCCGCATGCGGGC
CCTGAGCCAGGAGATCACCCGCGACTTCAACCTCCTGCAGGTCTCGGAGCCCTCGGAGCCAT
GTGTGAGATACCTGCCCAGGCTGTACCTGGACATACACAATTACTGTGTGCTGGACAAGCTG
CGGGACTTTGTGGCCTCGCCCCCGTGTTGGAAAGTGGCCCAGGTAGATTCCCTTGAAGGACAA
AGCACGGAAGCTGTACACCATCATGAACTCGTTCTGCAGGAGAGATTTGGTATTCCTGTTGG
ATGACTGCAATGCCTTGGAATACCCAATCCCAGTGACTACGGTCCTGCCAGATCGTCAGCGC
TAAGGGAACTGAGACCAGAGAAAGAACCCAAGAGAACTAAAGTTATGTCAGCTACCCAGACT
TAATGGGCCAGAGCCATGACCCTCACAGGTCTTGTGTTAGTTGTATCTGAAACTGTTATGTA
TCTCTCTACCTTCTGGAAAACAGGGCTGGTATTCCTACCCAGGAACCTCCTTTGAGCATAGA
GTTAGCAACCATGCTTCTCATTCCCTTGACTCATGTCTTGCCAGGATGGTTAGATACACAGC
ATGTTGATTTGGTCACTAAAAAGAAGAAAAGGACTAACAAGCTTCACTTTTATGAACAATA
TTTTGAGAACATGCACAATAGTATGTTTTTATTACTGGTTTAATGGAGTAATGGTACTTTTA
TTCTTTCTTGATAGAAACCTGCTTACATTTAACCAAGCTTCTATTATGCCTTTTTTCTAACAC
AGACTTTCTTCACTGTCTTTTCATTTTAAAAAGAAATTAATGCTCTTAAGATATATATTTTACG
TAGTGCTGACAGGACCCACTCTTTCATTGAAAGGTGATGAAAATCAAATAAAGAATCTCTTC
ACATGGA

1000 900 800 700 600 500 400 300 200 100 0

FIGURE 34

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA93011

><subunit 1 of 1, 136 aa, 1 stop

><MW: 15577, pI: 8.88, NX(S/T): 0

MRTPGPLPVLLLLLAGAPAARPTPPTCYSRMRALSQEITRDFNLLQVSEPSEPCVRYLPRLY
LDIHNYCVLDKLRDFVASPPCWKVAQVDSLKDKARKLYTIMNSFCRRDLVFLLDDCNALEYF
IPVTTVLPDRQR

Important features of the protein:

Signal peptide:

amino acids 1-19

Tyrosine kinase phosphorylation site.

amino acids 60-69

N-myristoylation site.

amino acids 16-22

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
214

FIGURE 35

GTCTCCGCGTCACAGGAACTTCAGCACCCACAGGGCGGACAGCGCTCCCCTCTACCTGGAGA
CTTGACTCCCGCGCGCCCCAACCTGCTTATCCCTTGACCGTCGAGTGTCAGAGATCCTGCA
GCCGCCAGTCCCGGCCCTCTCCCGCCCCACACCCACCTCCTGGCTCTTCCTGTTTTTAC
TCCTCCTTTTCATTCATAACAAAAGCTACAGCTCCAGGAGCCCAGCGCCGGGCTGTGACCCA
AGCCGAGCGTGGAAGAATGGGGTTCTCTCGGGACCGGCACTTGGATTCTGGTGTTAGTGCTCC
CGATTCAAGCTTTCCCCAACCTGGAGGAAGCCAAGACAAATCTCTACATAATAGAGAATTA
AGTGCAGAAAGACCTTTGAATGAACAGATTGCTGAAGCAGAAGAAGACAAGATTAAAAAAC
ATATCCTCCAGAAAACAAGCCAGGTCAGAGCAACTATTCTTTTGTGATAACTTGAACCTGC
TAAAGGCAATAACAGAAAAGGAAAAAATTGAGAAAGAAAGACAATCTATAAGAAGCTCCCCA
CTTGATAATAAGTTGAATGTGGAAGATGTTGATTCAACCAAGAATCGAAAACCTGATCGATGA
TTATGACTCTACTAAGAGTGGATTGGATCATAAATTTCAAGATGATCCAGATGGTCTTCATC
AACTAGACGGGACTCCTTTAACCGCTGAAGACATTGTCCATAAAATCGCTGCCAGGATTTAT
GAAGAAAATGACAGAGCCGTGTTTGACAAGATTGTTTCTAACTACTTAATCTCGGCCTTAT
CACAGAAAGCCAAGCACATACACTGGAAGATGAAGTAGCAGAGGTTTTACAAAAATTAATCT
CAAAGGAAGCCAACAATTATGAGGAGGATCCCAATAAGCCCACAAGCTGGACTGAGAATCAG
GCTGGAAAAATACCAGAGAAAGTGACTCCAATGGCAGCAATTCAAGATGGTCTTGCTAAGGG
AGAAAACGATGAAACAGTATCTAACACATTAACCTTGACAAATGGCTTGGAAGGAGAACTA
AAACCTACAGTGAAGACAACCTTTGAGGAACTCCAATATTTCCCAAATTTCTATGCGCTACTG
AAAAGTATTGATTGAGAAAAAGAAGCAAAAGAGAAAGAAACACTGATTACTATCATGAAAAC
ACTGATTGACTTTGTGAAGATGATGGTGAAATATGGAACAATATCTCCAGAAGAAGGTGTTT
CCTACCTTGAAAACCTTGGATGAAATGATTGCTCTTCAGACCAAAAACAAGCTAGAAAAAAT
GCTACTGACAATATAAGCAAGCTTTTCCCAGCACCATCAGAGAAGAGTCATGAAGAAACAGA
CAGTACCAAGGAAGAAGCAGCTAAGATGGAAAAGGAATATGGAAGCTTGAAGGATTCCACAA
AAGATGATAACTCCAACCCAGGAGGAAAGACAGATGAACCCAAAGGAAAAACAGAAGCCTAT
TTGGAAGCCATCAGAAAAAATATTGAATGGTTGAAGAAACATGACAAAAAGGGAAATAAAGA
AGATTATGACCTTTCAAAGATGAGAGACTTCATCAATAAACAAGCTGATGCTTATGTGGAGA
AAGGCATCCTTGACAAGGAAGAAGCCGAGGCCATCAAGCGCATTTATAGCAGCCTGTAAAA
TGGCAAAAGATCCAGGAGTCTTTCAACTGTTTCAGAAAACATAATATAGCTTAAACACTTC
TAATTCTGTGATTAAATTTTTTTGACCCAAGGGTTATTAGAAAGTGCTGAATTTACAGTAGT
TAACCTTTTACAAGTGGTTAAAACATAGCTTTCTTCCCGTAAAAACTATCTGAAAGTAAAGT
TGTATGTAAGCTGAAAAAAAAAAAAAAAAAAAAA

FIGURE 36

MGFLGTGTWILVLVLP IQAFPKPGGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPPEN
KPGQSNYSFVDNLNLLKAITEKEKIEKERQSIRSSPLDNKLNVEDVDSTKNRKLIDDYDSTK
SGLDHKFQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGLITESQA
HTLEDEVAEVLQKLISKEANNYEEDPNKPTSWTENQAGKIPKVT PMAAIQDGLAKGENDET
VSNTLTLTNGLERRTKTYSEDNFEELQYFPNFYALLKSIDSEKEAKEKETLITIMKTLIDFV
KMMVKYGTISP EEGVSYLENLDEMIALQTKNKLEKNATDNISKLPAPSEKSHEETDSTKEE
AAKMEKEYGSLKDSTKDDNSNPGGKTDEPKGKTEAYLEAIRKNIEWLKKHDKKGNKEDYDLS
KMRDFINKQADAYVEKGILDKEEA EAIKRIYSSL

N-glycosylation sites:

amino acids 68-71, 346-349, 350-353

Casein kinase II phosphorylation site:

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-
220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-
341, 369-372, 382-385, 386-389, 387-390

N-myristoylation sites:

amino acids 143-148, 239-244

1990-1991	1992-1993	1994-1995	1996-1997	1998-1999	2000-2001	2002-2003	2004-2005	2006-2007	2008-2009	2010-2011	2012-2013	2014-2015	2016-2017	2018-2019	2020-2021	2022-2023	2024-2025	2026-2027	2028-2029	2030-2031	2032-2033	2034-2035	2036-2037	2038-2039	2040-2041	2042-2043	2044-2045	2046-2047	2048-2049	2050-2051	2052-2053	2054-2055	2056-2057	2058-2059	2060-2061	2062-2063	2064-2065	2066-2067	2068-2069	2070-2071	2072-2073	2074-2075	2076-2077	2078-2079	2080-2081	2082-2083	2084-2085	2086-2087	2088-2089	2090-2091	2092-2093	2094-2095	2096-2097	2098-2099	2100-2101	2102-2103	2104-2105	2106-2107	2108-2109	2110-2111	2112-2113	2114-2115	2116-2117	2118-2119	2120-2121	2122-2123	2124-2125	2126-2127	2128-2129	2130-2131	2132-2133	2134-2135	2136-2137	2138-2139	2140-2141	2142-2143	2144-2145	2146-2147	2148-2149	2150-2151	2152-2153	2154-2155	2156-2157	2158-2159	2160-2161	2162-2163	2164-2165	2166-2167	2168-2169	2170-2171	2172-2173	2174-2175	2176-2177	2178-2179	2180-2181	2182-2183	2184-2185	2186-2187	2188-2189	2190-2191	2192-2193	2194-2195	2196-2197	2198-2199	2200-2201	2202-2203	2204-2205	2206-2207	2208-2209	2210-2211	2212-2213	2214-2215	2216-2217	2218-2219	2220-2221	2222-2223	2224-2225	2226-2227	2228-2229	2230-2231	2232-2233	2234-2235	2236-2237	2238-2239	2240-2241	2242-2243	2244-2245	2246-2247	2248-2249	2250-2251	2252-2253	2254-2255	2256-2257	2258-2259	2260-2261	2262-2263	2264-2265	2266-2267	2268-2269	2270-2271	2272-2273	2274-2275	2276-2277	2278-2279	2280-2281	2282-2283	2284-2285	2286-2287	2288-2289	2290-2291	2292-2293	2294-2295	2296-2297	2298-2299	2300-2301	2302-2303	2304-2305	2306-2307	2308-2309	2310-2311	2312-2313	2314-2315	2316-2317	2318-2319	2320-2321	2322-2323	2324-2325	2326-2327	2328-2329	2330-2331	2332-2333	2334-2335	2336-2337	2338-2339	2340-2341	2342-2343	2344-2345	2346-2347	2348-2349	2350-2351	2352-2353	2354-2355	2356-2357	2358-2359	2360-2361	2362-2363	2364-2365	2366-2367	2368-2369	2370-2371	2372-2373	2374-2375	2376-2377	2378-2379	2380-2381	2382-2383	2384-2385	2386-2387	2388-2389	2390-2391	2392-2393	2394-2395	2396-2397	2398-2399	2400-2401	2402-2403	2404-2405	2406-2407	2408-2409	2410-2411	2412-2413	2414-2415	2416-2417	2418-2419	2420-2421	2422-2423	2424-2425	2426-2427	2428-2429	2430-2431	2432-2433	2434-2435	2436-2437	2438-2439	2440-2441	2442-2443	2444-2445	2446-2447	2448-2449	2450-2451	2452-2453	2454-2455	2456-2457	2458-2459	2460-2461	2462-2463	2464-2465	2466-2467	2468-2469	2470-2471	2472-2473	2474-2475	2476-2477	2478-2479	2480-2481	2482-2483	2484-2485	2486-2487	2488-2489	2490-2491	2492-2493	2494-2495	2496-2497	2498-2499	2500-2501	2502-2503	2504-2505	2506-2507	2508-2509	2510-2511	2512-2513	2514-2515	2516-2517	2518-2519	2520-2521	2522-2523	2524-2525	2526-2527	2528-2529	2530-2531	2532-2533	2534-
-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-------

[illegible]

FIGURE 38

MALPPGPAALRHTLLLLPALLSSGWGELEPQIDGQTWAERALRENERHAFTCRVAGGPGTPR
LAWYLDGQLQEASTSRLLSVGGEAFSGGTSTFTVTAHRAQHELNCSLQDPRSGRSANASVIL
NVQFKPEIAQVGAKYQEAQGPGLLVVLFALVRANPPANVTWIDQDGPVTVNTSDFLVLDAQN
YPWLTNHTVQLQLRSLAHNLSVVATNDVGVTASLPAPGSPSRHPSLISSDSNNLKLNNVRLP
RENMSLPSNLQLNDLTPDSRAVKPADRQMAQNNSRPELLDPEPGGLLTSQGFIRLPVLGYIY
RVSSSVSSDEIWL

N-glycosylation sites:

amino acids 106-110, 119-123, 162-166, 175-179, 192-196, 205-209,
251-255, 280-284

Glycosaminoglycan attachment site:

amino acids 23-27

Casein kinase II phosphorylation sites:

amino acids 36-40, 108-112, 164-168, 282-286, 316-320

N-myristoylation sites:

amino acids 34-40, 89-95, 215-221, 292-298, 293-299

FIGURE 39

CGGGGACGGAAGCGGCCCCCTGGGCCCCGAGGGGCTGGAGCCGGGCGGGGCG**ATGT**GGAGCGC
GGGCCGCGGCGGGGCTGCCTGGCCGGTGCTGTTGGGGCTGCTGCTGGCGCTGTTAGTGCCGG
GCGGTGGTGCCGCCAAGACCGGTGCGGAGCTCGTGACCTGCGGGTCGGTGCTGAAGCTGCTC
AATACGCACCACCGCGTGCGGCTGCACTCGCACGACATCAAATACGGATCCGGCAGCGGCCA
GCAATCGGTGACCGGCGTAGAGGCGTCGGACGACGCCAATAGCTACTGGCGGATCCGCGGGC
GCTCGGAGGGCGGGTGCCCCGCGGGGTCCCCGGTGCGCTGCGGGCAGGCGGTGAGGCTCACG
CATGTGCTTACGGGCAAGAACCTGCACACGCACCACTTCCCGTCGCCGCTGTCCAACAACCA
GGAGGTGAGTGCCCTTTGGGGAAGACGGCGAGGGCGACGACCTGGACCTATGGACAGTGCGCT
GCTCTGGACAGCACTGGGAGCGTGAGGCTGCTGTGCGCTTCCAGCATGTGGGCACCTCTGTG
TTCCTGTCAGTCACGGGTGAGCAGTATGGAAGCCCCATCCGTGGGCAGCATGAGGTCCACGG
CATGCCCAGTGCCAACACGCACAATACGTGGAAGGCCATGGAAGGCATCTTCATCAAGCCTA
GTGTGGAGCCCTCTGCAGGTCACGATGAACTC**TGAG**TGTGTGGATGGATGGGTGGATGGAGG
GTGGCAGGTGGGGCGTCTGCAGGGCCACTCTTGGCAGAGACTTTGGGTTTGTAGGGGTCCTC
AAGTGCCTTTGTGATTAAAGAATGTTGGTCTATGAAA

FIGURE 40

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96857

><subunit 1 of 1, 221 aa, 1 stop

><MW: 23598, pI: 6.96, NX(S/T): 0

MWSAGRGGAAWPFVLLGLLLALLVPGGGAAKTGAELVTCGSVLKLLNTHHRVRLHSHDIKYGS
GSGQQSVTGVEASDDANSYWIRIRGGSEGGCPRGSPVRRCGQAVRLTHVLTGKNLHTHHFPSPL
SNNQEVSAFGEDGEGEDDLDLWTVRCGQHWEREAAVRFQHVGTSVFLSVTGEQYGSPIRGQH
EVHGMPSANTHNTWKAMEGIFIKPSVEPSAGHDEL

Important features of the protein:

Signal peptide:

amino acids 1-28

Glycosaminoglycan attachment site.

amino acids 62-66

N-myristoylation sites.

amino acids 16-22, 25-31, 27-33, 61-67, 71-77, 86-92, 87-93,
91-97, 190-196

Endoplasmic reticulum targeting sequence.

amino acids 218-223

FIGURE 41

GTTGCTATGTTGCCCAGGCTGGTCTTGAAGTGCCTTGACCTCCTAAAGTGTTGGAACCACAG
ACGTGAGCCACTCCACCCAGCCTAAAACCTTCATCTTCTTTGGATGAGATGAACACTTTTAAC
AAGAGAACAGGACTCTATATAAATCGCTGTGGGCTCACCACCTCTAAGGAGGAGCACTGACT
GAAGACAGAAAAATTGATGAACTGAAGAAGACATGGTCCATTATGCCTTACAACTTACACA
GTGCTTTGGGAATTCCAAAGTACTCAGTGGAGAGAGGTGTTTCAGGAGCCGTAGAGCCAGAT
CGTCATC**ATG**TCTGCATTGTGGCTGCTGCTGGGCCTCCTTGCCCTGATGGACTTGTCTGAAA
GCAGCAACTGGGGATGCTATGGAAACATCCAAAGCCTGGACACCCCTGGAGCATCTTGTGGG
ATTGGAAGACGTCACGGCCTGAACTACTGTGGAGTTCGTGCTTCTGAAAGGCTGGCTGAAAT
AGACATGCCATACCTCCTGAAATATCAACCCATGATGCAAACCATTTGGCCAAAAGTACTGCA
TGGATCCTGCCGTGATCGCTGGTGTCTTGTCCAGGAAGTCTCCCGGTGACAAAATTCTGGTC
AACATGGGCGATAGGACTAGCATGGTGCAGGACCCTGGCTCTCAAGCTCCCACATCCTGGAT
TAGTGAGTCTCAGGTTTCCCAGACAACTGAAGTTCTGACTACTAGAATCAAAGAAATCCAGA
GGAGGTTTCCAACCTGGACCCCTGACCAGTACCTGAGAGGTGGACTCTGTGCCTACAGTGGG
GGTGCTGGCTATGTCCGAAGCAGCCAGGACCTGAGCTGTGACTTCTGCAATGATGTCCTTGC
ACGAGCCAAGTACCTCAAGAGACATGGCTTC**TAA**CATCTCAGATGAAACCCAAGACCATGAT
CACATATGCAGCCTCAAATGTTACACAGATAAACTAGCCAAGGGCACCTGTAAGTGGGAAT
CTGAGTTTGACCTAAAAGTCATTAAAATAACATGAATCCCATTAACCAAAAAAAAAAAAAA

1000 900 800 700 600 500 400 300 200 100 0

FIGURE 42

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96867

><subunit 1 of 1, 194 aa, 1 stop

><MW: 21431, pI: 8.57, NX(S/T): 0

MSALWLLLGLLALMDLSESSNWGCYGNIQSLDTPGASCGIGRRHGLNYCGVRASERLAEIDM
PYLLKYQPMMQTIGQKYCMDPAVIAGVLSRKSPGDKILVNMGDRTSMVQDPGSQAPTSWISE
SQVSQTTEVLTTRIKEIQRRFPTWTPDQYLRGGLCAYSGGAGYVRSSQDLSCDFCNDVLARA
KYLKRHGF

Important features of the protein:

Signal peptide:

amino acids 1-19

N-myristoylation sites.

amino acids 23-29, 26-32, 35-41, 45-51, 50-56, 76-82, 156-162

Amidation site.

amino acids 40-44

FIGURE 43

TTGAAAATCTACTCTATCAGCTGCTGTGGTTGCCACCATTCTCAGGACCCTCGCCATGAAAG
CCCTTATGCTGCTCACCCTGTCTGTTCTGCTCTGCTGGGTCTCAGCTGACATTGCTGTGCTCAC
TCCTGCTACAAGGTCCCTGTGCTGGGCTGTGTGGACCGGCAGTCCTGCCGCCTGGAGCCAGG
ACAGCAATGCCTGACAACACATGCATACCTTGGTAAGATGTGGGTTTTCTCCAATCTGCGCT
GTGGCACACCAGAAGAGCCCTGTCAGGAGGCCTTCAACCAAACCAACCGCAAGCTGGGTCTG
ACATATAACACCACCTGCTGCAACAAGGACAAGTCAACAGCGCAGGACCCCGGCCCACTCC
AGCCCTGGGCCTTGTCTTCCTTACCTCCTTGGCTGGCCTTGGCCTCTGGCTGCTGCACTGAG
ACTCATTCCATTGGCTGCCCCCTCCTCCACCTGCCTTGGCCTGAGCCTCTCTCCCTGTGTCT
CTGTATCCCCTGGCTTTACAGAATCGTCTCTCCCTAGCTCCCATTTCTTTAATTAAACACTG
TTCCGAGTGGTCTCCTCATCCATCCTTCCCACCTCACACCCTTCACTCTCCTTTTTCTGGGT
CCCTTCCCACCTCCTTCCAGGACCTCCATTGGCTCCTAGAAGGGCTCCCCACTTTGCTTCCT
ATACTCTGCTGTCCCCCTACTTGAGGAGGGATTGGGATCTGGGCCTGAAATGGGGCTTCTGTG
TTGTCCCCAGTGAAGGCTCCCACAAGGACCTGATGACCTCACTGTACAGAGCTGACTCCCCA
AACCAGGCTCCCATATGTACCCCATCCCCCATACTCACCTCTTTCCATTTTGAGTAATAAA
TGTCTGAGTCTGGAAAAAAAAAAAAAAAAAAAA

1000
900
800
700
600
500
400
300
200
100
0

FIGURE 44

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96878

><subunit 1 of 1, 125 aa, 1 stop

><MW: 13821, pI: 8.60, NX(S/T): 2

MKALMLLTLSVLLCWVSADIRCHSCYKVPVLGCVDRQSCRLEPGQQCLTTHAYLGKMVFSN
LRCGTPEEPCQEAFNQTNRKLGITYNTTCNKDNCNSAGPRPTPALGLVFLTSLAGLGLWLLH

Important features of the protein:

Signal peptide:

amino acids 1-18

N-glycosylation sites.

amino acids 77-81, 88-92

N-myristoylation site.

amino acids 84-90

Ly-6 / u-PAR domain protein signature.

amino acids 85-98

FIGURE 45

ACGGGCCGCAGCGGCAGTGACGTAGGGTTGGCGCACGGATCCGTTGCGGCTGCAGCTCTGCA
GTCGGGGCCGTTTCCTTCGCCGCCGCCAGGGGTAGCGGTGTAGCTGCGCAGCGTCGCGCGCGCT
ACCGCACCCAGGTTTCGGCCCGTAGGCGTCTGGCAGCCCGGCGCCATCTTCATCGAGCGCCAT
GGCCGCAGCCTGCGGGCCGGGAGCGGCCGGGTACTGCTTGCTCCTCGGCTTGCATTTGTTC
TGCTGACCGCGGGCCCTGCCCTGGGCTGGAACGACCCTGACAGAATGTTGCTGCGGGATGTA
AAAGCTCTTACCCTCCACTATGACCGCTATAACACCTCCCGCAGGCTGGATCCCATCCCACA
GTTGAAATGTGTTGGAGGCACAGCTGGTTGTGATTCTTATACCCCAAAGTCATACAGTGTC
AGAACAAGGCTGGGATGGGTATGATGTACAGTGGAATGTAAGACGGACTTAGATATTGCA
TACAAATTTGGAAAACTGTGGTGAGCTGTGAAGGCTATGAGTCCTCTGAAGACCAGTATGT
ACTAAGAGGTTCTTGTGGCTTGGAGTATAATTTAGATTATACAGAACTTGGCCTGCAGAAAC
TGAAGGAGTCTGGAAAGCAGCACGGCTTTGCCTCTTTCTCTGATTATTATTATAAGTGGTCC
TCGGCGGATTCTGTAAACATGAGTGGATTGATTACCATCGTGGTACTCCTTGGGATCGCCTT
TGTAAGTCTATAAGCTGTTCTGAGTGACGGGCAGTATTCTCCTCCACCGTACTCTGAGTATC
CTCCATTTTCCCACCGTTACCAGAGATTCACCAACTCAGCAGGACCTCCTCCCCAGGCTTT
AAGTCTGAGTTCACAGGACCACAGAATACTGGCCATGGTGCAACTTCTGGTTTTTGGCAGTGC
TTTTACAGGACAACAAGGATATGAAAATTCAGGACCAGGGTTCTGGACAGGCTTGGGAACTG
GTGGAATACTAGGATATTTGTTTGGCAGCAATAGAGCGGCAACACCCTTCTCAGACTCGTGG
TACTACCCGTCCTATCCTCCCTCCTACCCTGGCACGTGGAATAGGGCTTACTCACCCCTTCA
TGGAGGCTCGGGCAGCTATTCGGTATGTTCAAACCTCAGACACGAAAACCAGAACTGCATCAG
GATATGGTGGTACCAGGAGACGATTAAAGTAGAAAGTTGGAGTCAAACACTGGATGCAGAAAT
TTTGGATTTTTTCATCACTTTCTCTTTAGAAAAAAAGTACTACCTGTTAACAATTGGGAAAAG
GGGATATTCAAAGTTCTGTGGTGTATGTCCAGTGTAGCTTTTTGTATTCTATTATTTGAG
GCTAAAAGTTGATGTGTGACAAAATACTTATGTGTTGTATGTCAGTGTAACATGCAGATGTA
TATTGCAGTTTTTTGAAAGTGATCATTACTGTGGAATGCTAAAAATACATTAATTTCTAAAC
CTGTGATGCCCTAAGAAGCATTAAAGAATGAAGGTGTTGTACTAATAGAACTAAGTACAGAA
AATTTAGTTTTAGGTGGTTGTAGCTGATGAGTTATTACCTCATAGAGACTATAATATTCTA
TTTGGTATTATATTATTTGATGTTTGCTGTTCTTCAAACATTTAAATCAAGCTTTGGACTAA
TTATGCTAATTTGTGAGTTCTGATCACTTTTGAGCTCTGAAGCTTTGAATCATTAGTGGTG
GAGATGGCCTTCTGGTAACTGAATATTACCTTCTGTAGGAAAAGGTGGAAAATAAGCATCTA
GAAGGTTGTTGTGAATGACTCTGTGCTGGCAAAAATGCTTGAAACCTCTATATTTCTTTTCGT
TCATAAGAGGTAAAGGTCAAATTTTTCAACAAAAGTCTTTTAATAACAAAAGCATGCAGTTC
TCTGTGAAATCTCAAATATTGTTGTAATAGTCTGTTTCAATCTTAAAAAGAATCA

FIGURE 46

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96889

><subunit 1 of 1, 339 aa, 1 stop

><MW: 36975, pI: 7.85, NX(S/T): 1

MAAACGPGAAGYCLLLGLHLFLLTAGPALGWNDPDRMLLRDVKALTLHYDRYTTSRRLDPIIP
QLKCVGGTAGCDSYTPKVIQCQNKGWDGYDVQWECKTDLDIAYKFGKTVVSCEGYESSEDQY
VLRGSCGLEYNLDYTELGQLKESGKQHGFAFSFYKSSADSCNMSGELITIVVLLGIA
FVVYKLFLSDGQYSPPPYSEYPPFSHRYQRFTNSAGPPPPGFKSEFTGPQNTGHGATSGFGS
AFTGQQGYENSGPGFWTGLGTGGILGYLFGSNRAATPFSDSWYYPSYPPSYPGTWNRAYSPL
HGGSGSYSVCSNSDTKTRTASGYGGTRRR

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 171-190

N-glycosylation site.

amino acids 172-176

Glycosaminoglycan attachment sites.

amino acids 244-248, 259-263, 331-335

Tyrosine kinase phosphorylation site.

amino acids 98-106

N-myristoylation sites.

amino acids 68-74, 69-75, 131-137, 241-247, 247-253, 266-272,
270-276, 278-284, 312-318